

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.



**An in depth study of human papillomavirus
diversity in South African women infected
with HIV**

ANNA T SALIMO

**A thesis submitted to the Faculty of Science in fulfilment of the
requirements for the degree of Master of Science in the
Department of Molecular & Cell Biology, University of Cape
Town**

November 2009



Name:	Anna Tina Salimo
Student Number	mnhann002
Course	MCB 5005W

Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

2. I have used the author and date convention for citation and referencing. Each contribution to this project from the work(s) of other people has been attributed, and has been cited and referenced.

3. This project is my own work.

4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature ASalimo

Date 24/05/2010

Table of Contents

ACKNOWLEDGEMENTS.....	iv
ABBREVIATIONS.....	v
ABSTRACT.....	vi
 CHAPTER 1 LITERATURE REVIEW.....	 2
 CHAPTER 2 GENOTYPING OF HPV FROM CERVICAL SPECIMENS FROM HIV INFECTED WOMEN.....	 28
 CHAPTER 3 ANALYSIS OF HPV 16 VARIANTS.....	 50
 CHAPTER 4 AN HPV-SPECIFIC METAGENOMIC ANALYSIS ON ONE HIV INFECTED SPECIMEN.....	 65
 CHAPTER 5 PCR SCREENING OF CLINICAL SPECIMENS FOR ADDITIONAL HPV TYPES.....	 90
 CHAPTER 6 CONCLUSION.....	 99
 APPENDIX A.....	 102
APPENDIX B.....	106
APPENDIX C.....	108
REFERENCES.....	109

Acknowledgements

I would like to express my sincere gratitude to:

My supervisors, Dr I. I. Hitzeroth, Prof A-L Williamson and Prof E. P. Rybicki for their advice, guidance and support throughout this work. They gave their time and listened willingly. They also showed immense patience during the course of this work and for that I am truly thankful.

Dr J. R. Moodley for kindly providing the cervical samples from HIV positive women in this study.

Mr Bruce Allan for the technical help, my laboratory colleagues, both in MCB and IIDMM, for helping with moral support and cheerfulness, as well as being pleasant to work with.

My family, especially my sweetheart, Luckmore Salimo, for all the support, encouragement and belief they showed in me, and for the sacrifices they made for me. All my friends who always supported, encouraged and cheered me up throughout this work. I am truly grateful.

Poliomyelitis Research Foundation, National Research Foundation and Margaret McNamara Memorial Fund for the financial support.

God Almighty, who made all this possible by his grace.

Abbreviations

μl	microlitre
A	Asian
AA	Asian-American
AC	Adenocarcinoma
Af	African
ASCUS	atypical squamous cells of unknown significance
ATPase	Adenosine Triphosphatase
bp	basepairs
BPCV1	Bandicoot papillomatosis carcinomatosis virus type 1
CIN	cervical intraepithelial lesions
DNA	Deoxyribonucleic acid
dNTP	deoxribonucleotriphosphate
E	European
GRE	glucocorticoid regulatory element
GS	Genome Sequencing
HC	Hybrid Capture
HIV	Human immunodeficiency virus
HPV	Human papillomavirus
HSIL	high-grade intraepithelial lesions
HSV	Herpes simplex virus
Kb	kilobases
LA	Luria Agar
LB	Luria-Bertani broth
LCR	long control region
LiPA	Line probe assay
LSIL	low-grade intraepithelial lesions
ml	milliliter
mRNA	messenger Ribonucleic acid
NRE	negative response element
ORF	open reading frame
PCR	Polymerase Chain Reaction
pRB	retinoblastoma
PRE	progesterone responsive element
PVs	papillomaviruses
PyVs	polyomaviruses
RCA	Rolling-circle amplification
RCR	Rolling-circle replication
SCC	squamous cell carcinoma
SHPC	streptavidin-horseradish peroxidase conjugate
STI	sexually transmitted infection
VLPs	virus-like particles
YY1	Yin-yan factor 1

Abstract

Cervical cancer is the second most common cancer affecting women and in most developing countries it remains the leading cause of cancer deaths. In South Africa, more than 3 400 women succumb to the disease every year and 1 in 31 women develop cervical cancer. The causative agent for cervical cancer is the Human papillomavirus (HPV). High-risk (carcinogenic) HPV types have been linked with 99% of the incidences of cervical cancer. The most common types identified in almost 70% of cervical cancer cases worldwide are HPV 16 and 18. HPV infection is very common in young healthy women and most immunocompetent individuals can clear HPV infection. However, in immunosuppressed women, clearance by host immune system is impaired. In addition, multiple HPV infections are quite common in women with Human immunodeficiency virus (HIV) infections.

The objectives of this study were to identify HPV types in South African women who also had HIV infection, and secondarily, to determine if recombination of HPV genomes occurs. Determining the HPV types circulating in this country is important to enable identification of most common HPV types, in order to guide the development of vaccines against HPV infection. HPV genotyping was performed by the commercial Roche Linear Array HPV Genotyping Test. The most common HPV types in South African women with HIV co-infection were HPV 61 (24%, 26/109), HPV 66 (18%, 20/109), HPV 58 (17%, 19/109), HPV 53 (17%, 18/109), HPV 45 (16%, 17/109), HPV 18 (16%, 17/109) and HPV 70 (16%, 17/109). Rolling circle amplification was performed on one specimen prior to pyrosequencing to detect more HPV types. Pyrosequencing detected HPV types 16, 30, 39, 40, 56, 74 and 86. It was observed that HPV types other than those included in the commercial test were present. These were HPV types 30, 74 and 86. Type-specific PCR was used to determine the prevalence of these types among all the women in this study. HPV 30 prevalence was 15% (16/109), the prevalence of HPV 74 was found to be 13% (14/109) and HPV 86 prevalence was only 5% (5/109).

The full length HPV 39 genome was analysed to detect recombination by primer-walking and there were no recombination events identified. However, point mutations were observed in the E1, E2/E4, E5 and L1 regions of HPV 39 genome.

This study also analysed the variants of HPV 16 among HIV infected women. HPV 16 is the most studied HPV genome worldwide and sequence variations within the genome led to identification of variants. Some of the HPV 16 variants have been claimed to be more often associated with disease than others; hence identification of variants can potentially predict risk of cervical disease and cancer. In this study, variants of HPV 16 were determined within the long control region (LCR). The variants of HPV 16; European, African type 1 and 2, Asian and Asian-American, were identified among 15 South African women. The most common variant was the African type 2 (60%, 9/15) and infection with more than one variant was common.

Chapter 1

Literature Review

1	INTRODUCTION.....	2
1.1	TRANSMISSION AND SYMPTOMS OF HPV INFECTION	3
1.2	HPV STRUCTURE.....	6
1.3	HPV REPLICATION.....	8
1.4	TRANSFORMATION AND PROGRESSION TO CANCER.....	10
1.5	EPIDEMIOLOGY OF HPV	12
1.5.1	<i>Classification</i>	12
1.5.2	<i>Global HPV prevalence</i>	15
1.5.3	<i>HPV types in HSIL, LSIL and ASCUS</i>	16
1.6	RISK FACTORS FOR HPV INFECTION	18
1.6.1	<i>Environmental factors</i>	18
1.6.2	<i>Host factors</i>	20
1.6.3	<i>Viral factors</i>	20
1.7	PREVENTION OF INFECTION AND TREATMENT OF HPV-ASSOCIATED DISEASE	23
1.8	OBJECTIVES OF THIS STUDY.....	26

1 Introduction

Papillomaviruses (PVs) belong to the family *Papillomaviridae*. This taxonomic family is divided into genus, species, types, subtypes and variants (de Villiers *et al.*, 2004). Different PVs are named according to the species they infect as they are species-specific. For example, Bovine PVs infect mostly cattle, cottontail rabbit PVs infect rabbits and human PVs infect humans (Howley, 1996). The discovery of PVs dates back to the early 1900s when their link to skin warts/papillomas in humans was discovered. Their association with cancer was not apparent until experimental work on rabbit papillomavirus during the 1930s provided the indication that PVs can cause cancer (Shah and Howley, 1996). Further experimentation led to the development of testing methods for Human papillomavirus (HPV) DNA. This major breakthrough in the early 1980s made it possible to establish the role of HPV in cervical cancer (Kjaer *et al.*, 1996). Since then, HPVs have been recognised as the causative agents for cervical cancer: HPVs have been detected in 99.7% of all cervical cancer cases (Walboomers *et al.*, 1999). Although it is a preventable disease, cervical cancer continues to be the second most common cancer in women worldwide (Kjaer *et al.*, 1996; Walboomers *et al.*, 1999).

HPVs consist of a diverse group of viruses that infect the cutaneous area and mucosal epithelia (Schiffman and Kjaer, 2003). HPV infection can result in papilloma formation, precancerous lesions and cervical cancer (section 1.2). Cervical cancer is not the only HPV induced cancer. Other types of human cancers such as skin, anal, head and neck cancer, vulvar/vaginal cancer in women and penile cancer in men have been reported to be HPV linked (Schiffman and Kjaer, 2003; Hoory *et al.*, 2008). More than 200 HPV types exist, of which more than 100 HPV full genomes are known (Lizano *et al.*, 2009). However, only a few have been classified into distinct categories: HPVs are classified into high-risk, low-risk and probable high-risk types based on epidemiological and phylogenetic grouping (Muñoz *et al.*, 2003). ‘Risk’ refers to the oncogenic potential of the virus types, therefore defining their link to cervical neoplasia and cervical cancer. The high-risk HPV types are strongly associated with the development of malignant lesions and cervical cancer. Low-risk HPV types are associated with benign lesions as they are associated with a low risk for development of cancerous lesions (Hoory *et al.*, 2008). The

probable high-risk types have been found in few cases of cervical cancer and there is not enough evidence to support their carcinogenicity (Schiffman *et al.*, 2009). Although HPV infection can cause cancer, not all high-risk HPV infections result in cervical cancer. Among the high-risk types, HPV 16 and 18 are the major causes of cervical cancer. HPV 16 and 18 account for about 70% of cervical cancer cases globally (Castle *et al.*, 2008). Often, other cofactors are involved in the carcinogenesis of HPV; these are discussed in section 1.6.

The burden of cervical cancer is more profound in Africa than in economically developed countries, due largely to the inadequate screening programmes for the detection of cervical lesions (Chirenje, 2005, Moodley *et al.*, 2006). In South Africa, cervical cancer is a major public health challenge as it is the second most common cancer in the country. According to the National Cancer registry, the lifetime risk of developing cervical cancer in 2001 was 1 in 31 women (Mqoqi *et al.*, 2004). There is an increased burden of cervical cancer due to the high prevalence of Human immunodeficiency virus (HIV) among women (Moodley *et al.*, 2006). The sub-Saharan African region has the highest HIV infected population in the world, with 67% of global HIV infections found in this region alone. The global statistics on HIV infection indicate that 50% of HIV infected people are women. In the sub-Saharan region alone, almost 60% of the HIV infected population are women (UNAIDS, 2008). This further worsens the plight of women who also acquire HPV infection. The impact of HIV in HPV-related disease is discussed in section 1.6.

In this review, the transmission of HPV, global epidemiology, risk factors for cervical cancer and the virus replication will be discussed. HPV-linked transformation of the cervix will also be discussed and some of the treatment and prevention strategies for HPV infection.

1.1 Transmission and symptoms of HPV infection

The specific modes of acquisition of HPV are not well understood but it is presumed that infection of actively proliferating basal epithelium through micro-lesions is the point of entry for the virus (Schiffman and Kjaer, 2003). Transmission of anogenital HPV types is mainly through sexual contact with an infected person (Kjaer *et al.*, 1996). Non-sexual transmission of HPV has been shown in perinatal infection, close non-sexual contact and

possibly through indirect transmission via fomites (Wheeler, 2002). Fomites are objects that can transmit infectious organisms (virus/bacteria) from one person to another. Infection by HPV through the non-sexual transmission is minimal. The modes of transmission of HPV among children are controversial but it has been reported that mother-to-child transmission can lead to respiratory papillomatosis (Hoory *et al.*, 2008). Some HPV types are acquired as commensals since they are part of the human flora and can be isolated from the skin of healthy individuals (Antonsson and Hansson, 2002).

HPV infects only the stratified squamous epithelial cells of the skin or the anogenital and oropharyngeal mucosal membranes in humans (zur Hausen, 2000). The virus was shown to be highly species-specific and has been implicated in a wide range of clinical manifestations among humans. Most HPV infections are, however, asymptomatic. In cases where symptoms are apparent, HPV infection causes warts and lesions in various areas (hands, feet, oral cavity and genitals), depending on the HPV type and site of infection (Hoory *et al.*, 2008). A few examples of HPV types and the lesions they cause are listed in Table 1.1. An HPV infection can spontaneously disappear after about 6-12 months, until a new HPV type infects the individual (Schiffman and Kjaer, 2003). The clearance of HPV infection is initiated by the cell mediated immune system of the host (Frazer, 2009). Persistent infection with high-risk HPVs, in combination with certain cofactors, can lead to development of pre-malignant lesions as shown in Figure 1.1 (de Cremoux *et al.*, 2003). The period from initial infection to invasive cancer may take between 10-20 years, during which a series of events occur (Trottier and Franco, 2006; Hoory *et al.*, 2008). Figure 1.1 illustrates the factors and path linked to the carcinogenesis of HPV infection.

Table 1.1: HPV infection and symptoms associated with certain HPV types

HPV type	Risk ^a	Site of infection	Symptoms/type of lesions
1, 60	—	cutaneous	Plantar warts
5 ^b , 8 ^b	—	cutaneous	Epidermodysplasia verruciformis
6 ^c , 11 ^c	low	mucosa	genital warts
16, 18	high	mucosa	Head and neck SCC
16, 18, 31, 33, 45 ^d	high	mucosa	genital warts, CIN, SCC

^a HPV types have been classified into high- and low-risk based on association with cervical cancer, most of the HPVs whose association with cancer is unknown/not yet investigated are not classified (refer to section 1.5.1). ^b HPV types associated with malignant epidermodysplasia verruciformis lesions. ^c In rare occasions, have been associated with carcinomas. ^d Other high-risk HPV types (Table 1.3) also cause similar symptoms. Cervical intraepithelial neoplasia (CIN), Squamous cell carcinoma (SCC). Head and neck carcinomas include cancers of the oral cavity, pharynx and larynx (modified from IARC, 2007; Muñoz *et al.*, 2006)

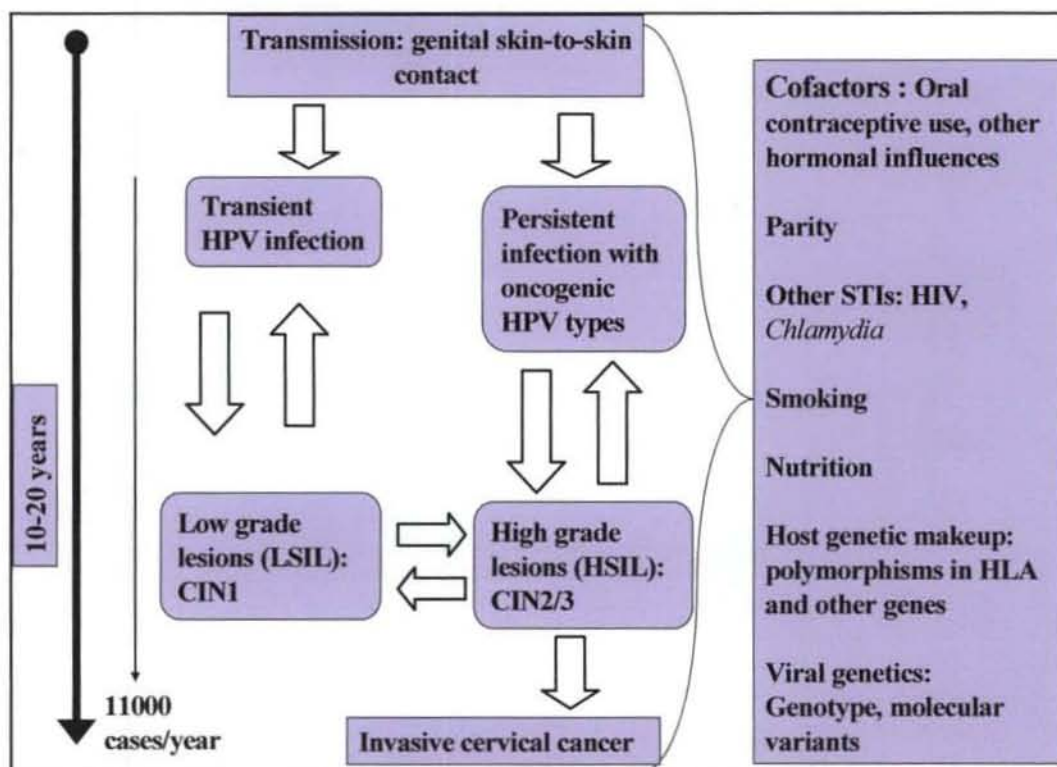


Figure 1.1: The progression to cervical carcinoma from the onset of HPV infection and the risk-factors for development of cervical disease. The cases/year refer to USA statistics. Most HPV infections clear spontaneously and a significant number of low-grade squamous intraepithelial lesions (LSIL) regress to normal while only a few high-grade squamous intraepithelial lesions (HSIL) spontaneously regress. cervical intraepithelial neoplasia (CIN), human leukocyte antigen (HLA) (Modified from Trottier and Franco, 2006; Hoory *et al.*, 2008).

The cervical lesions resulting from HPV infection have varying severity. These are cytologically determined as atypical squamous cells of unknown significance (ASCUS), low-grade squamous intraepithelial lesions (LSIL) and high-grade squamous intraepithelial lesions (HSIL). The LSIL encompass the early stage of cervical disease, classified histologically as cervical intraepithelial neoplasia 1 (CIN1). HSIL indicates the advanced stages of disease, histologically classified as cervical intraepithelial neoplasia 2/3 (CIN2/3) (Wright, 2007). The prevalence of different HPV types in these conditions is discussed in section 1.5.3. These conditions refer to squamous-cell carcinoma (SCC) conditions but glandular carcinoma can also occur. The glandular carcinomas are commonly referred to as adenocarcinoma (AC) (Howley and Lowy, 2007).

1.2 HPV structure

All HPVs have a similar genome organization with a total of 8 or 9 open reading frames (ORFs), depending on type. They have a double-stranded circular DNA genome of approximately 7900bp (Figure 1.2) (Bernard, 2002). The ORFs encode both the non-structural (early) and structural (late) proteins. The gene products are classified into early and late proteins based on functional action timing. The early proteins are transcribed from the early promoter and the late proteins are transcribed from the late promoter. The two structural late proteins of HPV, L1 and L2, assemble to form the viral capsid. The HPV capsid is non-enveloped particle with a T=7 icosahedral symmetry (Howley, 1996). The icosahedron is formed through the interaction of 72 pentamers of the major capsid protein L1. The minor capsid protein L2 associates with the L1 pentamers, and constitutes only one tenth of the mass of the viral capsid (Howley, 1996). The non-structural proteins comprise of a group of early proteins (E1, E2 and E4–E8) although not all HPVs have E5 (Schiffman *et al.*, 2005) and E8 genes. The E8 ORF is very short (12 amino acids) (Kadaja *et al.*, 2009), therefore it is usually omitted in structural diagrams of HPVs. The functions of the different proteins are summarised in Table 1.2. The HPV genome also contains a long control region (LCR), which contains regulatory elements for replication and transcription (Figure 1.2).

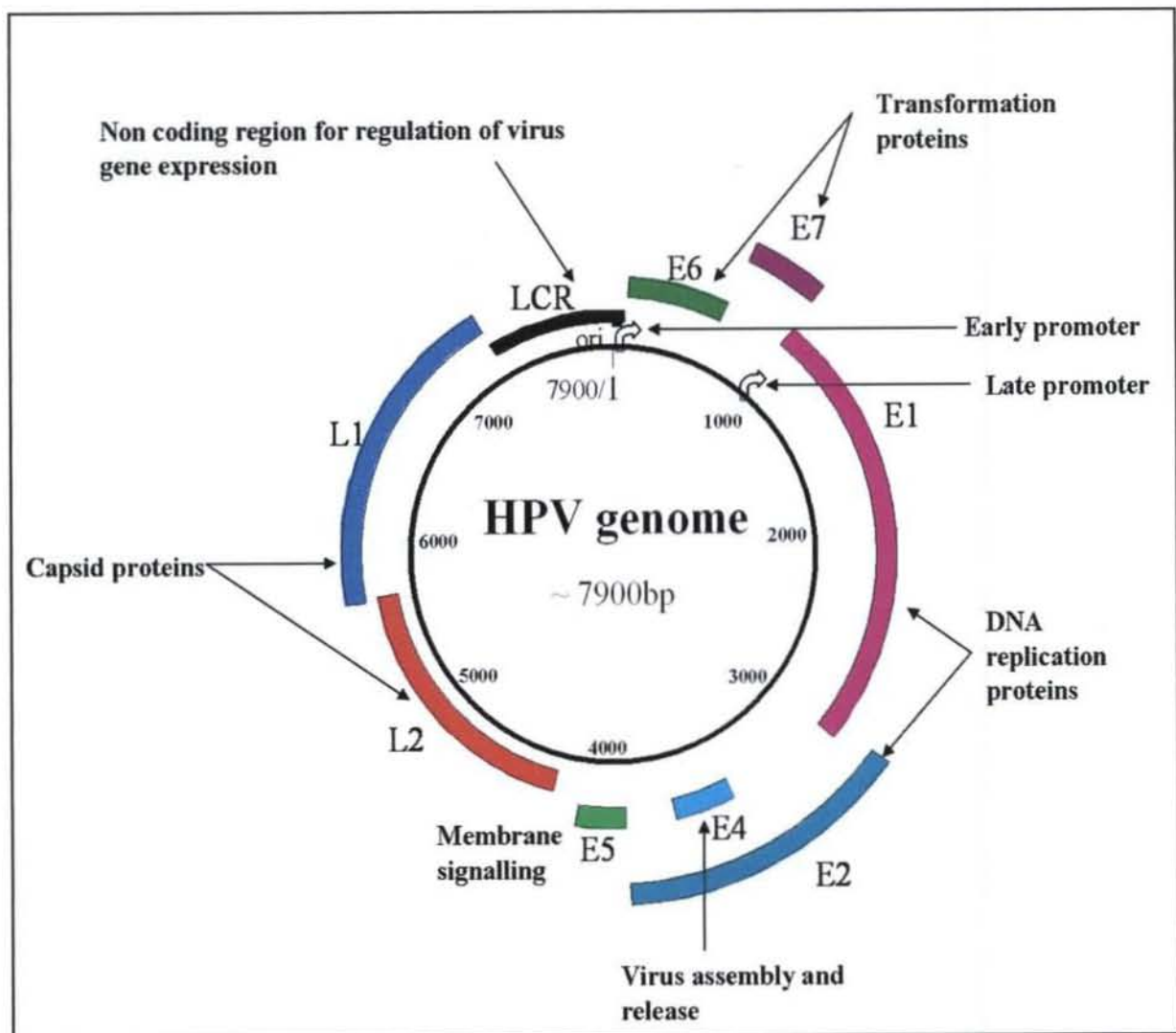


Figure 1.2: HPV DNA genome showing the different open reading frames (ORF). The E refers to early proteins and L refers to the late proteins. The short E8 ORF is omitted. The long control region (LCR), is a non-coding region whose role is in regulation of viral gene expression. The early promoter (p97 for HPV 16) is located next to the origin of replication (ori) site and the late promoter (p670 for HPV 16) is located further downstream of the early promoter. (Modified from Doorbar, 2006).

Table 1.2: Summary of the functions of HPV proteins.

HPV protein	Functional category	Primary role
E1	Replication	ATPase and DNA helicase that recognises and binds to viral origin of DNA replication, episomal maintenance, up-regulation of genome replication
E2	Replication, transcription	Episomal maintenance, repressor for E6 and E7, interacts with and recruits E1 to origin of replication, induces apoptosis
E4	Replication, transcription	Production of L2, disruption of cytokeratin filaments, induces G2 arrest, believed to facilitate virus assembly and release
E5	Replication, transcription, transformation	Stimulation of host growth factor receptors and protein kinases, induces unscheduled cell proliferation, inhibits apoptosis, inhibits traffic of major histocompatibility complexes to the cell surface
E6	Replication, transformation	Induces DNA synthesis, up-regulates telomerase, prevents cell differentiation, interacts with 4 classes of cellular proteins: transcriptional co-activators, proteins involved in cell motility, tumour suppressors and inducers of apoptosis; and replication and repair factors
E7	Replication, transcription, transformation	Up-regulates cell cycle, down-regulates Rb, induces unscheduled cell proliferation, interacts with histone acetyl transferases
E8	Replication	Down-regulation of episomal HPV DNA, long-term maintenance of HPV DNA
LI	Viral assembly	Major capsid protein, interacts with L2, interacts with cell receptors, encodes neutralising epitopes
L2	Viral assembly, Virion transport	Minor capsid protein, interacts with DNA and delivers it to replication sites, believed to facilitate virion assembly, encodes linear virus neutralising epitopes

(Thomison *et al.*, 2008 and IARC, 2007, Kadaja *et al.*, 2009)

1.3 HPV replication

The HPV life cycle is closely linked to the differentiation stages of host epithelial cells, and most of the life cycle is restricted to the epithelium (Kadaja *et al.*, 2009). After infection of the basal cells of the epithelia (Figure 1.3), viral DNA is transported to the nucleus for transcription and replication.

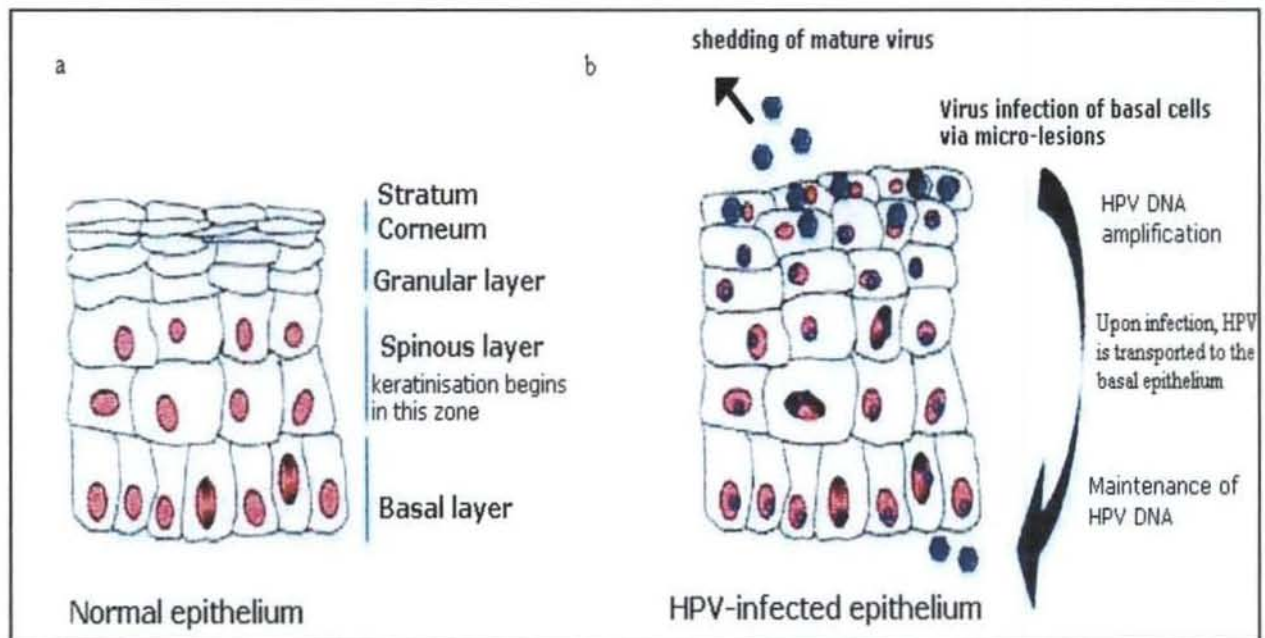


Figure 1.3: The infection cycle of HPV. a: The different layers of the normal epithelium are indicated. b: The HPV infected epithelium shows the level of replication in the different epithelium layers. HPV DNA replication is at low levels within the basal cells and is increased in the granular layer. To complete the life cycle, virion production, assembly and release of mature virus occur in the stratum corneum (Modified from Longworth and Laimins, 2004).

The mechanism of replication in PV genomes is determined by the state of the genome within the host cell and the stage of the virus life cycle (Kadaja *et al.*, 2009). In the undifferentiated epithelial cells, where non-productive HPV infection occurs, HPV DNA replication occurs bidirectionally via theta structures. This replication mechanism initiates replication at the origin of replication site under the regulation of E1 and E2 (Flores and Lambert, 1997; Kadaja *et al.*, 2009). However, upon differentiation of cells, the mechanism may switch to a unidirectional rolling-circle replication mechanism (Flores and Lambert, 1997), which is common in organisms with circular DNA genomes and plasmids (Rector *et al.*, 2004). The differentiated epithelial cells are characterised by productive HPV infections, and the viral genome is in the episomal form (Thomison *et al.*, 2008).

DNA replication of the Bovine papillomavirus 1 (BPV) has been extensively studied and has therefore provided the background for HPV DNA replication studies. After infection of the basal cells of the epithelia, the virus experiences three stages of replication (Thomison *et al.*, 2008). The first stage is the propagation of viral DNA to about 50-100 copies/cell and the second stage is the maintenance of latent infection at low copy

numbers in undifferentiated cells (Howley, 1996; Liu and Melendy, 2002). During the second stage the virus only replicates in synchrony with host cell cycle, while maintaining the episomal state. The final stage is the vegetative replication, which only occurs in terminally differentiated cells (Liu and Melendy, 2002; Kadaja *et al.*, 2009). Amplification of viral DNA during this phase may be supported by E4 (Thomison *et al.*, 2008). In this stage, control of copy number appears to be lost with a concomitant increase in the viral DNA copy number. Expression of the late proteins, L1 and L2, and new virus assembly occurs following vegetative replication. Virus particles are sloughed off with dead cells from the epidermis, completing the virus life cycle (Kadaja *et al.*, 2009).

Viral replication is regulated by the E1 and E2 proteins. The E1 protein enrolls cellular factors to allow viral DNA replication to proceed in the absence of host cell replication (Brown *et al.*, 2008; Thomison *et al.*, 2008). HPV E1 on its own has low DNA binding affinity, but it complexes with E2 and is targeted to the replication origin to initiate DNA replication (del Mar Pena and Laimins, 2002; Brown *et al.*, 2008). The levels of E1 and E2 increase upon cellular differentiation due to the activation of a late promoter, leading to an increase in the genome copy number (Thomison *et al.*, 2008). The increase in E2 proteins results in a regulatory effect on the transcription of the major transforming proteins E6 and E7 (section 1.4).

1.4 Transformation and progression to cancer

Transformation of host cells is a property of the high-risk HPV types. This is achieved through the oncoproteins E6 and E7. E5 is also considered a transforming protein although its role in transformation is limited. It has been demonstrated that E5 is not expressed in cells with the integrated HPV genome in cancer tumours, and therefore does not contribute to transformation (McLaughlin-Drubin and Münger, 2009). In those HPVs that express E5 protein, it enhances the immortalization potential of E6 and E7 (Münger and Howley, 2002; IARC, 2007). E6 and E7 oncoproteins stimulate cell-cycle progression by binding to cell-cycle regulators, p53 and retinoblastoma (Rb), respectively, which induces transformation of HPV-infected basal cells. E6 binds to p53 and stimulates its degradation through an ubiquitin-dependent protease system (Longworth and Laimins, 2004). Normal function of p53 protein is to arrest cells in G₁ phase following DNA damage during viral infection, thereby inducing apoptosis in the

affected cells. The blocking of apoptosis through degradation of p53 can lead to progression to malignancy (Longworth and Laimins, 2004). Expression of high-risk HPV E6 and E7 keeps the cells immortalised through their interaction with p53 and Rb proteins (McLaughlin-Drubin and Münger, 2009).

The E6 and E7 of low-risk HPVs are not capable of inducing transformation despite their ability to bind the cell-cycle regulators. This is due to a difference in the amino acid composition of the binding sites between the high- and low-risk HPVs (Longworth and Laimins, 2004). The binding of low-risk HPV E6 protein to p53 occurs in a manner that does not induce the degradation of p53. High-risk HPV E7 binds to Rb with increased efficiency than low-risk HPV E7, and targets it for degradation. Low-risk HPV E7 is unable to facilitate degradation of Rb due to decreased affinity in binding sites between the high- and low-risk HPVs (Münger and Howley, 2002; Longworth and Laimins, 2004).

The integration of HPV genome into host DNA by non-homologous recombination often disrupts the E1/E2 ORFs of viral genome leading to the loss of expression of E2 (Figure 1.3) (Jeon *et al.*, 1995; Kurvinen *et al.*, 2000). Integration occurs when E1, which is responsible for episomal genome maintenance, is insufficiently produced (Thomison *et al.*, 2008). When the genome is in episomal condition, the E2 protein regulates the expression of E6 and E7 proteins by suppressing the E6/E7 p97 gene promoter. This regulatory effect of E2 is lost upon integration, thus E6 and E7 transcription is up-regulated. This leads to an irreversible commitment by cells to enter S phase and replicate as cell-cycle regulators are disrupted by E6 and E7 (Longworth and Laimins, 2004). Progression to HSIL and cervical cancer may result due to increased E6 and E7 expression, which increases transformation of the cervix at the transformation zone. In LSIL, the high-risk HPV genomes exist as episomes while they are often integrated in HSIL and invasive cancer cases (Longworth and Laimins, 2004). This indicates that the progression of LSIL to HSIL is linked to integration of HPV DNA (IARC, 2007). HPV 16 DNA was found to be integrated in about 72% of cervical cancer cases. In contrast, HPV 18 was integrated in 100% of cases of cervical cancer, which possibly explains its association with more aggressive cancers (Tjalma *et al.*, 2005).

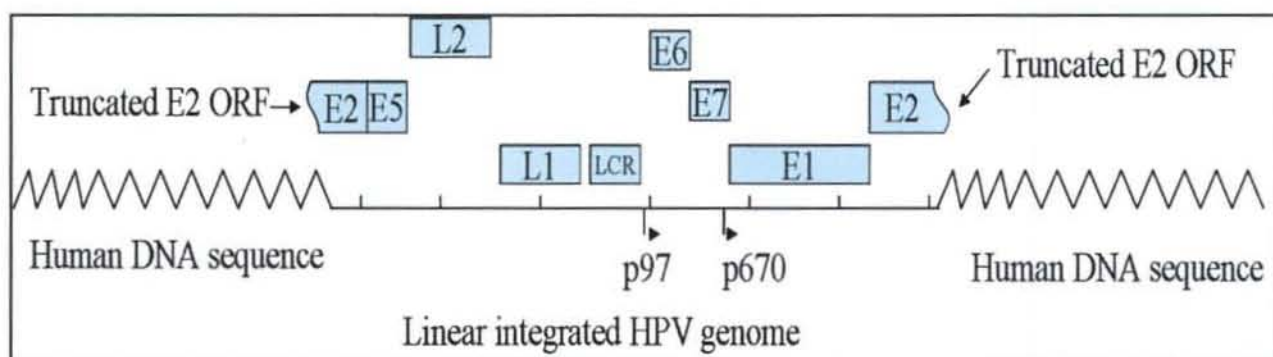


Figure 1.3: Integration of HPV DNA. HPV 16 DNA is integrated into the host DNA, resulting in an incomplete E2 ORF. The transcription of genes in the integrated genome is regulated by the same promoters (p97 and p670) as in the episomal condition. Open reading frame (ORF); Long control region (LCR) (Modified from Doorbar and Sterling, 2001).

1.5 Epidemiology of HPV

1.5.1 Classification

HPV types are grouped into two categories on the basis of the site of infection: these are the mucosal and cutaneous types (de Villiers, 1989). Further classification of HPV into high- and low-risk types is dependent on the epidemiological data for the specific types in cancers (Schiffman *et al.*, 2009). The high- and low- risk HPV types are grouped within the mucosal HPVs, while there are a lot more ‘unclassified’ types which are found within the cutaneous HPV types. Figure 1.4 shows the phylogenetic differentiation between cutaneous and mucosal HPV types.

Among the known HPV types, about 40 of them infect the genital tract. Eleven of these are consistently classified as high-risk (16, 18, 31, 33, 35, 39, 45, 51, 52, 56 and 58) (Munoz *et al.*, 2003; Bernard, 2005). These HPV types are associated with malignant lesions and invasive cancer, as shown in Table 1.1. From the study by Munoz and others (2003), 15 HPV types emerged as high-risk, based on their detection in prevalence studies. These included the 11 mentioned above, and HPV types 59, 68, 73, and 82 (Table 1.3). The number of putative high-risk types varies from about 13 to 19, which highlights the lack of clear-cut criteria for classifying HPV into high- and low-risk groups (Muñoz *et al.*, 2003). As an example to highlight this classification problem, HPV 73, which is phylogenetically related to HPV 34 (Figure 1.4), is grouped into high-risk category but HPV 34 remains ‘unclassified’. Also of significance is that some studies

classified HPV 73 as phylogenetically low-risk type, although it is generally grouped together with the high-risk types based on epidemiology studies (Muñoz *et al.*, 2003). This shows that although some HPV types may be closely related, their prevalence and correlation to cervical cancer within the population can be quite different. The impact of this is that classification of HPV may always be under review as more types are discovered or as definite evidence of some types in cervical cancer becomes available.

Table 1.3: Classification of HPV by cervical oncogenicity

Risk Classification	HPV types
High-risk	16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, 82
Probable high-risk	26, 53, 66
Low-risk	6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81, CP6108 (HPV 89)
Undetermined risk	34, 57, 83

Classification of HPV into high-, low- and probable high-risk types (Adapted from Baseman and Koutsky, 2005). This classification system will be referred to in this review and the rest of the document.

A recent review on the classification of some weakly carcinogenic HPV types showed that HPV 68 may be re-classified as a probable high-risk HPV (Schiffman *et al.*, 2009). This is in contrast to the high-risk classification shown in Table 1.3. In addition to the probable high-risk group, another classification category has been defined. This is the possibly high-risk group, which includes HPVs 26, 53, 66, 70, 73 and 82. According to the classification reviews, these HPV types are considered to be possible high-risk types since evidence for carcinogenicity is not sufficient (Schiffman *et al.*, 2009). In comparison to the classification in Table 1.3, it shows that HPV 82 and 73 are no longer considered as definite high-risk types. Other HPV types which were not studied in the IARC meta-analysis (HPV 30, 34, 67, 85 and 97) are phylogenetically related to some high-risk types (Figure 1.4, clade a). These are classified as possible high-risk types. This classification was based on the lack of sufficient evidence for carcinogenicity (Schiffman *et al.*, 2009). This re-classification of some HPV types clearly highlights the challenges faced when attempting to classify HPV types according to carcinogenicity.

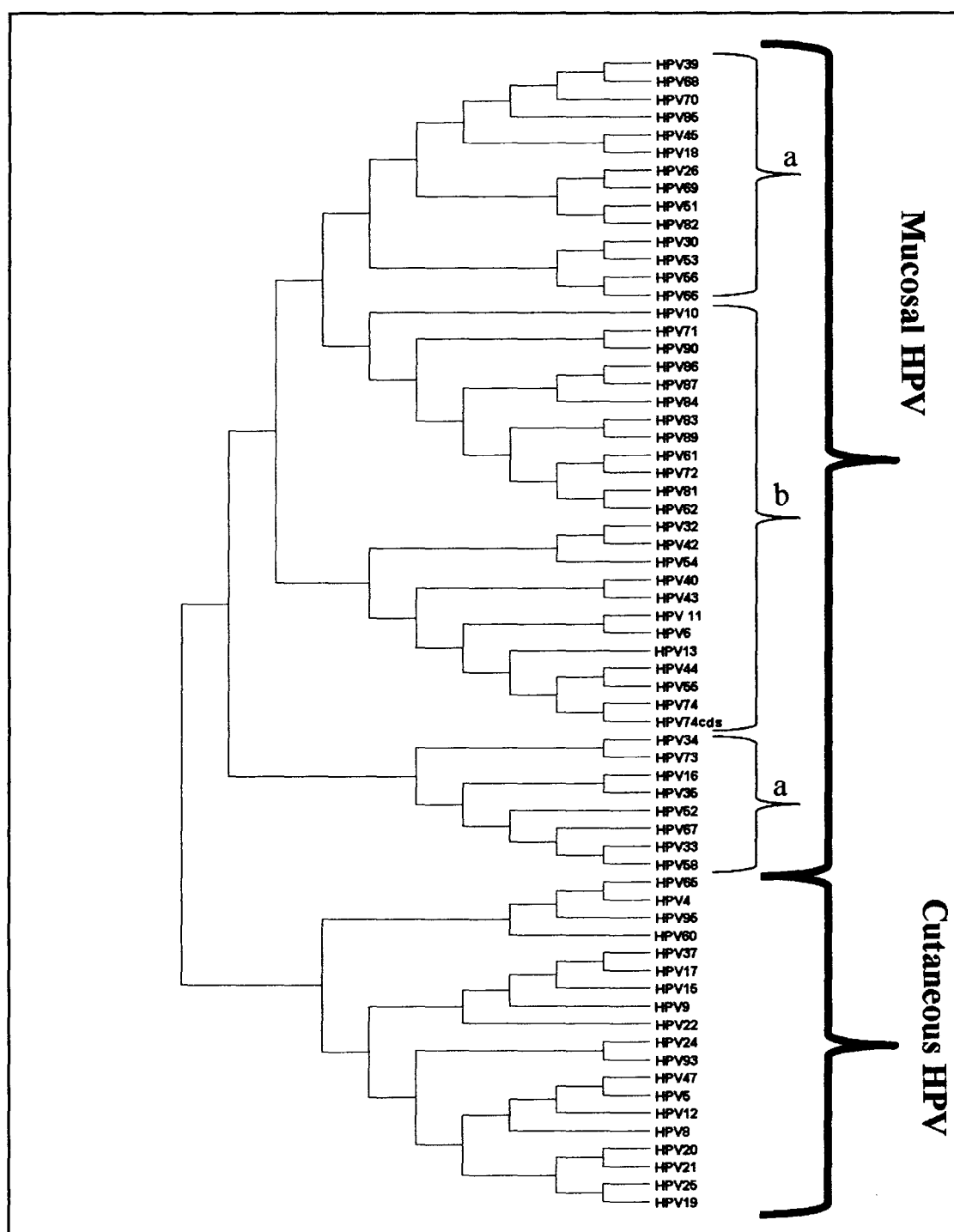


Figure 1.4: Evolutionary relationship of 65 HPV types classified into mucosal and cutaneous types. The full length HPV genome sequences were obtained from the GenBank database (NCBI, www.ncbi.nlm.nih.gov) and analysis was performed using MEGA4 program. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987) and the evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura *et al.*, 2004). Mucosal types are sub-divided into high- and low-risk, with clades a and b showing species with high- and low-risk HPV types, respectively. Some HPV types included in clade a (HPV 30, 69, 70 and 85) are not classified into the high-risk category although they are phylogenetically related to the high-risk or probable high-risk types. Phylogenetic analysis was conducted in MEGA4 (Tamura *et al.*, 2007).

1.5.2 Global HPV prevalence

The prevalence of HPV gives a measure of the percentage in a population of people who have new, persistent or recurring HPV infection. HPV can be detected in cytologically normal women (Clifford *et al.*, 2005) as well as in most cases of genital neoplasia and cancer (Clifford *et al.*, 2006a; Muñoz *et al.*, 2003). The distribution of HPV types is different according to region and also differs among women with normal Pap smears and those with cervical neoplasia and cancer (Clifford *et al.*, 2006a).

The most common types of HPV in women without cervical abnormalities worldwide, in decreasing order of frequency, are HPV 16, 42, 58, 31, 18, 56, 81, 35, 33, 45, and 52 (Clifford *et al.*, 2005). Sub-Saharan (Nigeria) women, however, are less likely to be infected with HPV 16 as compared to their counterparts in Europe, although they could be infected with other high- and low-risk HPV types (Clifford *et al.*, 2005). The likely reason for this difference seems to be a mutualistic effect between the biological properties of the virus and the host immune system (Clifford *et al.*, 2006a). HPV 35 is the most common high-risk type among sub-Saharan African (Nigeria) women without cervical abnormalities (Clifford *et al.*, 2005). In contrast to this, HPV 83 was found to be the most prevalent type in South African women without cervical abnormalities, followed by HPV 53 and lastly HPV 16 (Allan *et al.*, 2008). Clearly, this shows that the prevalence rates for the different HPV types are different when compared across countries. The burden of genital infection and cancers was found to be highest on the African continent compared to other regions. The genital tract infection among women with normal cytology in Africa is 21.3%, in America 14.5%, in Asia 10.9% and Europe has the least infections at 9.7% of the women surveyed (WHO/ICO Report, 2009).

The general distribution of HPV in women worldwide shows that HPV 16 is the most common type detected in both normal Pap smears and cervical cancers (Clifford *et al.*, 2005; Muñoz *et al.*, 2003). Among women with cervical cancer, Europe and North America have a higher than average prevalence of HPV 16 (Clifford *et al.*, 2006a). Muñoz and others (2003), in a 12 year study of cervical cancers in 9 countries on prevalence of HPV, also reported a worldwide prevalence of 58.9% for HPV 16 (Muñoz *et al.*, 2003). Their study tested 1918 women with cervical cancer and 1739 of them were positive for HPV DNA. From that study they also noted that HPV 18 was the second

most common type with an overall prevalence (among 1739 HPV infected women) of 15%; 3.7%, 4.4% and 27.9% in Spain, Colombia and Philippines, respectively (Muñoz *et al.*, 2003). The highest prevalence of HPV 45 was in the Philippines (15.7%) but it was very low in Spain and Colombia. HPV 52 was highest in Peru (8.6%) (Muñoz *et al.*, 2003) and HPV types 31 and 35 were found more commonly in Latin America (Muñoz *et al.*, 2003; Clifford *et al.*, 2006a). The eight most common HPV types detected in cervical cancer patients worldwide, in decreasing order of frequency, are HPV 16, 18, 45, 31, 33, 52, 58 and 35 (Muñoz *et al.*, 2003). HPV 16 alone accounts for about 56% of all cervical cancers (Barnabas *et al.*, 2006), but it was reported to be more prevalent in squamous cell carcinoma (SCC) cases than in adenocarcinoma (AC) cases (Clifford *et al.*, 2006a). The prevalence of different HPV types may vary between SCC and AC. For example, the occurrence of HPV 16 in women with SCC was shown to be much higher than that of HPV 18. In women with AC, however, HPV 18 was more prevalent than HPV 16 (Clifford *et al.*, 2006a).

Despite the differences in burden of HPV infection, it is clear that in all the regions there are HPV types that are shared between women without cervical abnormalities and those with cancer. These are mainly the high-risk types. They are, in no particular order, HPV types 16, 18, 45, 31, 33, 52, 58 and 35.

1.5.3 HPV types in HSIL, LSIL and ASCUS

HPV distribution differs significantly between lesions of different severity. There are HPV types that are most often found in a particular type of lesion. The HPV types found in women with HSIL are similar to those found in cervical cancer, with HPV 16 being the most common type. The eight most common HPV types in HSIL, in decreasing order of frequency, are 16, 31, 58, 33, 18, 33, 52, and 35. Surprisingly, HPV 18 seemed to be less prevalent in the HSIL across all regions, as it was the third or fourth most common. It is also striking that the related HPV 45 is absent in most regions except in South and Central America. North, South and Central America recorded HPV 18 as the third most common HPV detected in HSIL. In Africa, HPV 18 is the fourth most common (Clifford *et al.*, 2006a). This pattern was confirmed from a South African study involving 53 women with HSIL. In these women, HPV types 16 and 35 emerged as the predominant types, with HPV 18 further down the hierarchy of common HPV types in HSIL (Allan *et*

al., 2008). In addition, HPV 58 and 66 have also been identified among women with HSIL in South Africa (Said *et al.*, 2009). When comparing the HPV distribution in HSIL in South Africa to the global distribution, HPV 66 was only observed among South African women. HPV types 31, 33 and 45 were not among the common types in the South African general population. This reflects the differences that can be found in HPV distribution in different regions.

Of the 8308 HPV cases in a worldwide study by Clifford and colleagues (2006a), 64% were positive for LSIL. However, only 3% of the LSIL positive cases came from Africa, whereas most of the cases of HPV were from Europe and America (Clifford *et al.*, 2006a). Due to the small population size of African LSIL cases in that study, the HPV distribution in Africa may not be representative of some African countries. From that study, HPV 16 emerged as the most common type in all cases of LSIL, followed by HPV 31, then 51, 53, 56, 52, 18, 66, and 58 in decreasing order of frequency. Other HPV types were detected as well but only in about 5% of the LSIL cases, showing the wide heterogeneity of HPV in LSIL (Clifford *et al.*, 2006a). A different pattern of HPV distribution was observed in a study using samples obtained from South Africa only. In that study, Allan and others (2008) described HPV 52 as the prominent type, followed by types 53, 16 and 35 among 57 women with LSIL (Allan *et al.*, 2008). HPV 31 and 51 were not detected from these South African women while HPV 52, less frequent in Europe and America, was the most common type in women with LSIL. In these women, HPV 16 was the third most common unlike in European and American samples where it was the most frequent type. The distribution of HPV in HIV infected women, who were at higher risk of infection with more HPV types than HIV negative women, is discussed in section 1.6.3.1.

Few studies report on the HPV types observed in atypical squamous cells of unknown significance (ASCUS). One such study was conducted on South African samples, where women with ASCUS were most commonly infected with HPV 52, 45, 84, and 16, in order of decreasing frequency (Allan *et al.*, 2008).

The HPV distribution in the different types of lesions and cervical cancer highlights the HPV types that should be of concern in public health sector. In both LSIL and HSIL, only

the high-risk HPV types are more prevalent and this increases risk of lesions developing to cervical cancer.

1.6 Risk factors for HPV infection

The development to cervical cancer depends on other factors acting together with the oncogenic HPV types (de Cremoux *et al.*, 2003). In addition to HPV and other sexually transmitted infections (STIs including HIV, Herpes simplex virus type 2 (HSV-2), *Chlamydia trachomatis*), smoking, use of oral contraceptives, immune suppression, exposure to UV light, high parity, family history and multiple sexual partners have been implicated as cofactors for cervical cancer (Marais *et al.*, 2008; Hoory *et al.*, 2008; Castellsague and Munoz, 2003). These exogenous as well as endogenous factors influence progression to HSIL and cervical cancer by increasing damage to tissues (Castellsague and Muñoz, 2003; Marais *et al.*, 2008). Cofactors in cervical cancer can be divided into three main groups. These groups consist of environmental/exogenous cofactors, viral cofactors and host cofactors. The environmental factors include oral contraceptive use, tobacco smoking, diet, cervical trauma, co-infection with HIV and other STIs (Castellsague and Muñoz, 2003). Viral factors include co-infection with specific HPV types or variants, the viral load and viral integration into host DNA. The host cofactors like hormones, genetic factors and other host immune response factors also contribute to progression of cervical disease (Castellsague and Muñoz, 2003). These cofactors influence the acquisition of HPV infection; increase the risk of HPV persistence and increase the risk of progression from HPV infection to HSIL and invasive cancer (Castellsague and Muñoz, 2003). A few of these cofactors are discussed in the following sections.

1.6.1 Environmental factors

1.6.1.1 Oral contraceptives

The use of oral contraceptives has been associated mainly with AC *in situ*. With longer periods of oral contraceptive use, the risk of progression to cervical cancer is increased in the presence of HPV (Castellsague and Muñoz, 2003). There is no increased risk of cervical neoplasia for up to four years of oral contraceptive usage but for longer than five

years, there is a four-fold increased risk for invasive cervical cancer and three-fold risk for carcinoma *in situ* (CIS) (Castellsague and Muñoz, 2003).

Hormone related pathways may promote integration of HPV DNA into the host genome, although the mechanisms are not clearly understood. The integration of HPV DNA results in deregulation of oncogenic proteins E6 and E7 (Section 1.4) (Castellsague and Muñoz, 2003). Estradiol, an estrogen, was shown to possibly stimulate transcription of HPV 16 E6 and E7 within cell lines with integrated HPV 16 (Wheeler, 2002; Castellsague and Muñoz, 2003). It has been hypothesised that progesterone, which suppresses maturation of the cervical epithelium and retards keratinisation, may make the cervix more prone to tumour development (Shapiro *et al.*, 2003). Oral contraceptives might also facilitate HPV reactivation or persistence (Castellsague and Muñoz, 2003). In contrast to the claim that use of oral contraceptives may increase risk of HPV-induced cervical neoplasia and cancer, Shapiro and co-workers (2003) argued that there is no significant increase in risk (Shapiro *et al.*, 2003). In South Africa, screening of women for HPV infection is relatively uncommon but the use of oral contraceptives as well as injectible progesterone is high (Shapiro *et al.*, 2003). This potentially increases the number of women who develop cervical disease and cervical cancer. However, in a study in South Africa by Shapiro and co-workers (2003) on oral contraceptive use, there was no significant increase in the risk of cervical cancer with use of oral contraceptives although modest risk of cervical cancer could not be ruled out for oral contraceptive use for over 10 years or more (Shapiro *et al.*, 2003).

1.6.1.2 Parity

High parity has been found to be consistent with increased risk of both cervical cancer and CIS as most of the major studies among HPV positive women reported an increased risk of HSIL/cervical cancer with increasing number of pregnancies (Wheeler, 2002). High parity may increase the risk of cervical cancer because it maintains the transformation zone on the exocervix for many years, thereby increasing exposure to HPV and other cofactors. Giving birth at an early age, especially when one is already infected with HPV, has also been associated with increased risk of cervical cancer (Wheeler, 2002). Hormonal changes and immunodepression during pregnancy (high estrogen and progesterone levels) may modulate the immune response to HPV and / risk

of persistence (Castellsague and Muñoz, 2003). The same effect could also result from use of hormonal contraceptives, which also increase levels of progesterone and estrogen (Castellsague and Muñoz, 2003). Hormonal and immunological mechanisms have been hypothesized as plausible explanations for the association between parity and HSIL/cervical cancer among HPV infected women (Castellsague and Muñoz, 2003).

1.6.1.3 Smoking

Multiple studies conducted on the association of cervical cancer with smoking have consistently shown that tobacco smoking increases the risk of developing HSIL and cervical cancer (Wheeler, 2002). Malignant transformation of HPV 16-immortalised human endo-cervical cells due to cigarette smoke has been proven. Furthermore, nicotine and tobacco-specific carcinogens have been detected in cervical mucus of smokers (Castellsague and Muñoz, 2003; Wheeler, 2002). The inability of tobacco smokers to exert an effective immune response against viral infections may also increase the risk of developing cervical cancer (Castellsague and Muñoz, 2003).

1.6.2 Host factors

Some host factors that have been suggested to increase the risk of developing cancer include immunosuppression and the presence of genetic disorders like Fanconi anaemia. Individuals with this disease have a defective DNA repair system, thereby increasing the risk of cancer by more than 150 fold, 700 fold and 4000 fold for head, neck and vulvar cancer, respectively (Hoory *et al.*, 2008). Immunosuppression can be caused by several conditions including organ transplants and HIV infection. Only the latter is addressed in this review.

1.6.3 Viral factors

Although HPV plays a significant role in the development of cervical cancer, it has been shown that not all HPV types are responsible for cervical disease. This means that the development of cervical disease depends on the infecting HPV types. Some HPV types have different variants with different oncogenic potential, as has been predicted for some HPV 16 variants (Tu *et al.*, 2006). Co-infection with multiple HPV types, especially high-risk, is also associated with a high-risk for CIN. A high HPV viral load predicts an increased risk for HSIL or carcinoma, especially in HIV positive women whose viral load

is consistently high (Frisch *et al.*, 2000). Co-infection with other viruses like HIV and HSV may also aggravate outcome of HPV infection (Palefsky *et al.*, 1991). Only HPV infection in women with HIV is discussed below. As discussed in section 1.4, HPV integration plays a pivotal role in the transition from CIN to invasive carcinoma since the majority of cervical cancer cases have an integrated genome (Tjalma *et al.*, 2005). Therefore, the ability of an HPV type to integrate into the host DNA increases the risk of progression to cervical neoplasia and invasive cancer.

1.6.3.1 HIV and HPV

HPV infects the epithelial cells while HIV targets the lymphoid tissue cells (zur Hausen, 2000; Peterlin and Trono, 2003). The difference in the target cells for infection suggests that dual infection of same cell is unlikely. The mechanisms of HPV-HIV interaction are still poorly understood, although it is known that HIV *tat* protein stimulates replication of HPV through its effect on the LCR (Palefsky, 2006; Vernon *et al.*, 1993). In some studies on the relationship between HPV and HIV infection, it has been shown that interaction between HIV and HPV 16 leads to HPV persistence and cervical neoplasia (Palefsky and Holly, 2003). This is a result of gradual damage to the immune system and loss of immune control in HIV infected individuals, thus leading to the loss of the HPV specific immune response (Palefsky, 2006).

The gradual increase in cases of HPV induced cancer in developed countries may be a direct result of the HIV epidemic. HIV positive women show an increased risk of infection with HPV (Palefsky and Holly, 2003) compared to HIV negative women. They also experience faster progression of cervical abnormalities when compared to HIV negative women (Hawes *et al.*, 2003; Marais *et al.*, 2008). A high prevalence of CIN was detected in HIV positive adults, as well as a notably faster progression to high-grade lesions (Palefsky and Holly, 2003). There is also a relatively higher prevalence of cutaneous and anogenital warts, which also tend to be larger, more numerous and more recurrent in HIV infected individuals compared to HIV negative women (Harwood and Proby, 2001).

Due to immunosuppression in HIV positive women, there is an increased risk of infection by low-risk HPVs that could lead to more cervical disease than in immunocompetent

adults (Luque *et al.*, 1999, Clifford *et al.*, 2006b). The low immunity levels in HIV positive individuals seem to allow for HPV DNA viral replication, resulting in increased detection of HPV DNA (Palefsky and Holly, 2003). This is because the HIV *tat* protein seems to stimulate the HPV LCR, increasing binding of replication factors (Vernon *et al.*, 1993). This highlights the inverse relationship between HPV viral load and immune competence (Frisch *et al.*, 2000). Individuals with HIV are at an increased risk of having HPV induced HSIL when compared to HIV negative women, especially in individuals with high-risk HPV types. Low-risk HPV 11 and 61 were found in HIV infected women with HSIL and in whom there were no high-risk types detected (Hawes *et al.*, 2003; Clifford *et al.*, 2006b). This suggests that in HIV immunosuppressed women, the low-risk HPV types may lead to more severe lesions.

Multiple infections with HPV are consistent with loss of immune control (lower CD4+) in the HIV positive women hence the increased disease severity (Frisch *et al.*, 2000). HIV-induced immunosuppression impairs the cell-mediated immunity against HPV clearance. Therefore, the level of clearance of HPV in these individuals is very low, leading to the persistence of HPV infection (Frisch *et al.*, 2000).

The most commonly occurring cancers in HIV positive people were found to be vulvar/vaginal cancer and cervical cancer in women, penile cancer in men and invasive anal cancer in both men and women (Frisch *et al.*, 2000). In separate studies on women with HPV in Kenya, Zambia, Zimbabwe, Uganda and South Africa, the results indicated an increased risk of cervical disease for HIV positive women (Hawes *et al.*, 2003; Sahasrabuddhe *et al.*, 2007; Chirenje, 2005; Moodley *et al.*, 2006). Overall, this indicated the increased burden of cervical disease on the African continent. The prevalence of HPV in Africa increased to 56% in HIV positive women compared with 23% in HIV-negative women (Hoory *et al.*, 2008; Clifford *et al.*, 2006b). These prevalence rates were observed in women with normal cytology. This implies that among women with cervical abnormalities and cancer, HPV prevalence rates are more likely to be higher in HIV infected women.

In America, studies investigating the predominant HPV types in HIV positive individuals reported higher prevalence of HPV 68, 56, 53 and 16 compared to their HIV negative counterparts. HPV 11 and 61 were also found more commonly in HIV positive women

with HSIL than in HIV negative women (Luque *et al.*, 2006, Clifford *et al.*, 2006b). In studies done in Africa on women co-infected with HIV, the prevalent types in Ugandan women were HPV 52, 58 and 16 (Blossom *et al.*, 2007; Banura *et al.*, 2008). In Zambia, the most common HPV types, in decreasing order of frequency, were HPV 52, 61, 58, 53 81 and 16 (Sahasrabuddhe *et al.*, 2007). These studies, although in different continents (America and Africa), showed that HPV 16 was the third or fourth most common HPV type in HIV infected women. Clearly, this shows that in HIV infected individuals, there is a shift towards other HPV types in addition to HPV 16, which is otherwise the most common type globally. A major difference in HPV 16 distribution was noted in HSIL cases where the prevalence of HPV 16 in women with HIV was lower than that in HIV negative women (Clifford *et al.*, 2006b). In a South African study including 37 HIV positive women, HPV 45 and 66 were reported to be more common in HIV infected women. HPV 18 prevalence was lower in the HIV infected women compared to its prevalence in HIV negative women (Marais *et al.*, 2008). In comparison to HPV types in other continents, America for example, the distribution of HPV is different to that in African countries when considering prevalence rates. These differences in the distribution and prevalence patterns of HPV types in different countries necessitate the profiling of HPV types in a country. Establishing the diversity of HPV genotypes among HIV-infected individuals is important if the development of effective vaccines is to be successful in South Africa and the Sub-Saharan region.

1.7 Prevention of infection and treatment of HPV-associated disease

Currently there is no cure for HPV infection although lesions like warts can be cured. Most treatment forms available only target the warts and lesions (LSIL and HSIL) while a few may increase killing of the etiological agent (Stanley, 2003). Treatment options for most HPV-induced cutaneous warts include the use of cryosurgery with liquid nitrogen, hyfrecation and topically administered drugs (Stanley, 2003; Leman and Benton, 2000). Recurrence of warts may occur as the virus still persists. Genital lesions can be treated by use of topical applications of creams such as imiquimod and 5-Fluorouracil. Imiquimod induces a strong T-cell mediated immune response to kill the virus although it (HPV) may not completely be eradicated (Stanley, 2003). Although 100% efficacy cannot be guaranteed, genital warts and anogenital lesions can be treated by ablative methods, including surgical excision. Interferons are also used for the treatment of warts due to

their immunostimulatory effects but as with other methods, recurrence rates are high (Stanley, 2003). The screening of women for cervical disease (section 2.1) is useful to detect LSIL and HSIL (Naucler *et al.*, 2007) and treatment of these lesions is possible through cervical conisation to prevent progression to cervical cancer (Stanley, 2003). Despite the availability of several treatment options, recurrence of lesions may occur in future. Although LSIL cases can be treated by ablative methods, sometimes it is not treated but regular follow-up is recommended to monitor for progression to HSIL since the rate of regression of LSIL is high (Stanley, 2003).

The use of prophylactic vaccines against HPV is most likely to significantly reduce incidences of cervical cancer since it has been documented that 100% of cervical cancer cases have HPV. Prophylactic vaccination against HPVs 6, 11, 16 and 18 has drastically reduced persistent HPV infection and cervical disease incidences (Castle *et al.*, 2008). The two vaccines available (Gardasil produced by Merck pharmaceuticals and Cervarix manufactured by GlaxoSmithKline) protect against the types most associated with cervical cancer (HPVs 16 and 18). Gardasil also protects against types that cause genital warts (HPVs 6 and 11) (Herrero, 2009). The vaccines consist of virus-like particles (VLPs) made from L1 proteins, which self-assemble and present epitopes in a similar way to the virus without being infectious. Current production of the vaccines is in insect cells (Cervarix) and yeast cells (Gardasil). Based on phylogenetic similarities between some HPV types, the current vaccines may also protect against types 31 and 45, which are closely related to HPV 16 and 18, respectively (Herrero, 2009). Prophylactic vaccination before sexual activity therefore is an effective preventive measure against HPV infection. HPV infection can also be prevented by abstinence and faithfulness to uninfected partner, assuming the candidate is not already infected. Condom usage is also advised, although this only provides partial protection against HPV infection (Baseman and Koutsky, 2005).

The differential distribution of HPV types in cervical disease in a region or country has important implications for the application and possible future development of vaccines. Current prophylactic vaccines do not include some of the HPV types commonly observed in South African women with LSIL and HSIL who are infected with HIV. Prevention of cervical disease from a broader spectrum of HPVs infecting women would require vaccines that include more HPV types (polyvalent vaccines). The successful development

of these vaccines can be achieved when the most prevalent types in cervical cancer in a region or country are identified.

1.8 Objectives of this study

The distribution of carcinogenic HPV types is region-specific. The high and increasing HIV prevalence in South Africa may significantly change the spectrum of HPVs associated with cervical disease. The purpose of this study was therefore firstly, to investigate the diversity of HPV types in South African women who are HIV positive. Most studies investigating HPV types among women have been performed in other African countries and few of these were among HIV positive individuals. In this study, HPV genotyping was achieved by using the commercial Roche Linear Array HPV Genotyping test. Pyrosequencing was also used for HPV genotyping to confirm the results of the Roche Linear Array Genotyping test and identify HPV types that are not detected by the commercial genotyping test. The prevalence of the HPV types not detected by the commercial test was determined by using type-specific PCR.

HIV infected women are often infected with multiple HPV types and have higher viral load, providing an ecological niche for recombination between HPV types. The second objective was therefore to investigate the possibility of recombination among HPV genomes in this population of HIV-infected women. This was achieved by analysis of full-length HPV genome through primer walking sequencing.

The third and final objective was to determine the variants of HPV 16 that exist in women with HIV co-infection. This was achieved by amplification of the long control region and identification of single nucleotide changes within this region. Some of the variants of HPV 16 have been linked to increased rate of progression to disease and these have been isolated in other countries. In South Africa, the comparison between variants and disease severity has not been studied; therefore, we wanted to investigate the correlation between isolated variants and cervical disease severity in patients.

Chapter 2

Genotyping of HPV from cervical specimens from HIV infected women

2	INTRODUCTION	28
2.1	SCREENING FOR CERVICAL DISEASE	28
2.2	HPV DNA TESTING	29
2.2.1	<i>Digene Hybrid Capture test for HPV.....</i>	<i>29</i>
2.2.2	<i>HPV DNA testing by PCR.....</i>	<i>31</i>
2.2.3	<i>HPV distribution in South Africa.....</i>	<i>36</i>
2.3	MATERIALS AND METHODS	37
2.3.1	<i>Study population</i>	<i>37</i>
2.3.2	<i>DNA extraction</i>	<i>37</i>
2.3.3	<i>PCR amplification and HPV genotyping</i>	<i>38</i>
2.4	RESULTS	39
2.4.1	<i>Low-risk HPV types</i>	<i>43</i>
2.4.2	<i>High-risk HPV types</i>	<i>43</i>
2.5	DISCUSSION	45

2 Introduction

The long premalignant phase in HPV infections provides the opportunity to detect and treat early disease, thereby preventing the progression to cervical cancer (Barnabas *et al.*, 2006). The high- and low-risk HPV types that can cause disease have been identified and their detection in individuals is important for monitoring infected women for disease progression. Different assay methods have been developed to detect HPV types in epidemiological and clinical studies, with some methods distinguishing up to 37 types (Gravitt *et al.*, 1998, 2000; Clavel *et al.*, 1998; Kleter *et al.*, 1998, 1999). It is beneficial to identify the most common oncogenic types within populations in order to guide vaccine development research. HPV infection is mostly asymptomatic, but also manifests more frequently as malignant disease in infections with high-risk HPVs (de Villiers *et al.*, 2004). This also occurs more often in HIV positive women, hence the need to monitor these women for cervical lesions. Primary screening for cervical disease (section 2.1) has been and continues to be practiced in order to identify women at risk of developing cervical cancer. Although primary screening has been useful in detecting asymptomatic HPV infections, the sensitivity of diagnosis of CIN2/3 is increased by the use of HPV DNA testing (Cox, 2006).

2.1 *Screening for cervical disease*

Cervical cancer is a preventable disease in an environment where detection, treatment and follow-up facilities for HPV infection are adequate. The Papanicolaou staining test (Pap smear) is used as the primary screening method for cervical disease. This procedure detects precancerous lesions when used repeatedly (Naucner *et al.*, 2007), but lacks high sensitivity. It is also subjective to adequate sampling and owing to the repetitive nature of reading, results in greater interpretation errors. Further tests can be performed in women with abnormal cytology: this is achieved through colposcopy, biopsy and HPV DNA testing. Colposcopy is a technique that uses a colposcope to magnify the cervix in order to detect any abnormalities, whereas biopsy involves the removal of a small piece of tissue for diagnostic purposes (IARC, 2007). Although they have been widely used in the past, serological tests are unreliable in the detection of HPV infection, hence the development of more reliable testing methods that target viral DNA (Schiffman and

Kjaer, 2003). These technological advances include the Digene HPV Test Hybrid Capture 2 (HC2) and polymerase chain reaction (PCR) based methods (section 2.2.2).

2.2 HPV DNA testing

Anogenital HPV prevalence in young healthy women is high (Cox, 2006). Most infections are transient; however, women with abnormal Pap test results require follow-up, referral for further testing and treatment when diagnosed with CIN2/3 (Cox, 2006). In order to provide cost-effective follow-ups and treatment of CIN1, HPV DNA testing can be used to determine persistent HPV infections. HPV DNA testing increases the identification of women likely to progress to CIN2/3 who require close monitoring. A combination of Pap cytology and HPV DNA testing provides 96-100% sensitivity for the detection of CIN2/3 and cervical cancer (Cox, 2006). Therefore, detection of HPV DNA could be used alongside cytology methods to improve specificity for cervical disease (Söderlund-Strand *et al.*, 2005). This is recommended especially for women 30 years old and above. For the clarification of ASCUS, HPV DNA testing can be used together with cytology for women of any age (Cox, 2006). The HC2 assay for high-risk HPVs was approved by the FDA for cervical cancer screening, and has therefore been incorporated into screening programmes in the USA (Wright *et al.*, 2007).

2.2.1 Digene Hybrid Capture test for HPV

The HC2 assay is used in many clinical diagnostic laboratories as the non-PCR HPV DNA detection method. The second-generation version of the assay, HC2, has an increased sensitivity compared to the first-generation assay, HC1 assay (Poljak *et al.*, 1999). HC2 is a commercial enzyme-linked immunosorbent assay on cervical scrapings. The HC2 assay uses a cocktail of probes for 13 high-risk HPV types (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68) and five low-risk types (HPV 6, 11, 42, 43 and 44) (Clavel *et al.*, 1998). The assay is based on sandwich hybridisation using HPV RNA probes against HPV DNA. The RNA-DNA hybrids are captured onto the surface of microplates coated with monoclonal antibodies specific for RNA-DNA hybrids. This is followed by detection with a non-radioactive alkaline phosphatase- conjugated antibody catalysing the cleavage of a chemiluminescent substrate to emit light (Figure 2.1). A positive result is recorded when the relative light unit reading is equal to or greater than the mean values from positive controls (Poljak *et al.*, 1999; Clavel *et al.*, 1998). The

value of the positive control is generally regarded as 1, equivalent to 1pg/ml HPV DNA (5000 copies) (Poljak *et al.*, 1999). This method, however, can not be used to identify the specific HPV genotypes but only to confirm the presence of HPV DNA from the above mentioned HPV types (Muñoz *et al.*, 2003). HC2 tests for high- and low-risk HPV types are commercially available. The HC2 high-risk HPV Test, which detects the high-risk HPV types mentioned above, is the one most commonly used. This is because medical governing bodies like the American Society for Colposcopy and Cervical Pathology (ASCCP) and American College of Obstetricians and Gynecologists (ACOG), among others, recommend testing for high-risk HPV types only for clinical purposes. This is due to the strong correlation of cervical disease to the high-risk HPV types (Wright *et al.*, 2007).

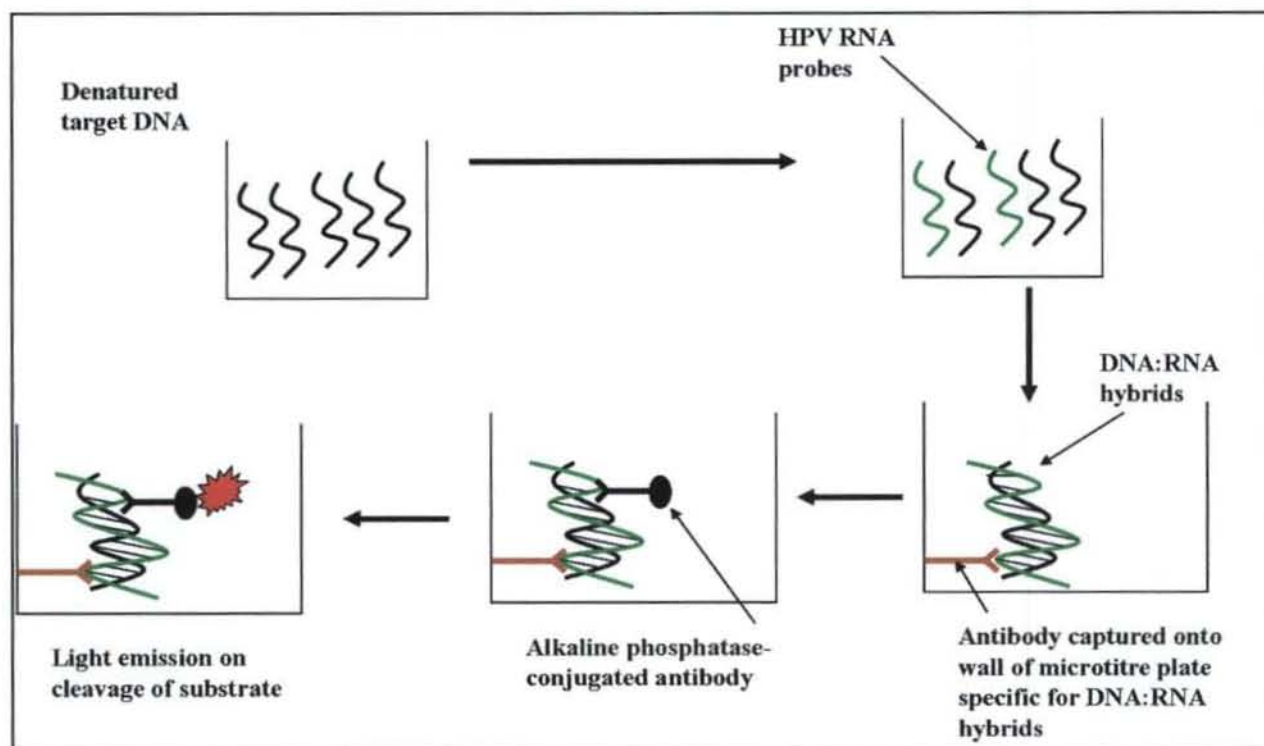


Figure 2.1: Schematic representation of the HC2 assay.

The use of HC2 gave false-negative results in some squamous cell carcinoma cases where high-risk HPV DNA was detected by PCR (Poljak *et al.*, 1999; Seme *et al.*, 2006). The success of HC2 may be influenced by the severity of lesions and the viral load in the specimen: very low levels of HPV DNA may give a negative result. Despite this, the level of specificity for disease and sensitivity for HPV DNA in HC2 assay is considered

adequate for use in cervical screening, and the specificity for disease is higher than that of PCR-based detection (Poljak *et al.*, 1999; Hesselink *et al.*, 2008). The sensitivity for HPV DNA of HC2 is comparable to that of PCR-based methods (Söderlund-Strand *et al.*, 2005).

2.2.2 HPV DNA testing by PCR

The majority of studies on HPV DNA detection have been done using PCR methods with consensus primers (Clifford *et al.*, 2006a). Many of these PCR methods used in HPV detection are 'in-house' developed, with a few which are commercially available. Most PCR based methods are followed by detection assays that allow for the identification of the specific HPV genotypes. These assay methods are mostly useful for detecting HPV in persistent infections; hence their ability to detect transient clinically irrelevant genotypes has to be as low as possible. This is critical in ensuring reliable results on specificity for disease and sensitivity for HPV DNA from high-risk persistent types responsible for high grade lesions (Hesselink *et al.*, 2008). Due to the occurrence of multiple infections in the population, the success of these methods is also evaluated on the ability to detect co-infection with different HPV genotypes. In clinical settings it is important to detect the persistent infections in order to avoid non cost-effective follow-ups in women with normal cytology (Hesselink *et al.*, 2008).

The major disadvantage of using PCR amplification is that there is competition for PCR primers when the samples have multiple HPV infections. This results in false-negative results in some cases (Lin *et al.*, 2008; Huang *et al.*, 2006). The sensitivity for HPV DNA of PCR detection methods varies greatly according to assay characteristics, type and quality of biological specimen and reagents. The sensitivity varies across the different 'in-house' PCR methods (discussed in sections 2.2.2.1-3). The differences in results obtained arise due to the amplicon sizes and type of specimen as well as quality of DNA sample (Clifford *et al.*, 2006a). It is also possible that the differences are due to primer design methods, which may contribute to the sensitivity for HPV DNA for each primer set. To reduce the differences in the sensitivity of these PCR methods, there is a need to introduce standardised primers, methods, reagents and possibly reference samples for an internationally approved uniform method for HPV detection. However, this has not been implemented despite the observed differences in the 'in-house' detection methods

(Clifford *et al.*, 2006a). Currently, there are two PCR based systems that are commercially available; The Linear Array HPV Genotyping Test (Roche Diagnostics, Germany) and the INNO-Line probe assay (Innogenetics, Belgium). However, there is no mandate to use these in all HPV genotyping work. As a result, other different 'in-house' methods are also used, creating differences in results. This lack of uniformity in PCR is a major drawback when comparing results in epidemiological studies. Some of the commonly used PCR methods are detailed below.

The primers used in PCR methods amplify a broad spectrum of HPV types in a single PCR reaction. Most HPV types are closely related; therefore primers have been designed to target the conserved L1 gene of HPV genomes (Wheeler, 2002). These primers include MY09/11, PGMY09/11, GP5+/6+ and SPF1/2 or SPF10, shown in Figure 2.2 (Manos *et al.*, 1989; Gravitt *et al.*, 2000; de Roda Husman *et al.*, 1995; Kleter *et al.*, 1999). The PGMY09/11 primer set is a modification of MY09/11 primers (Gravitt *et al.*, 2000). The modified primer set SPF1/GP6+ has also been used for detection of HPV DNA (Huang *et al.*, 2006). In Figure 2.2, the location of the various primer sets within the HPV genome with their respective PCR product sizes is shown.

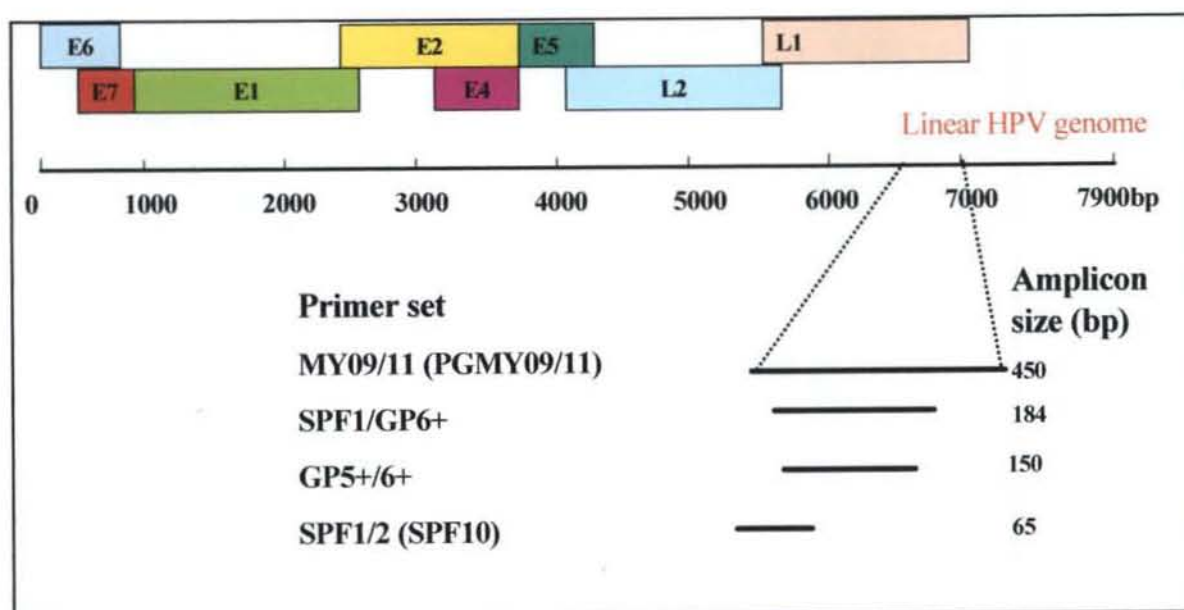


Figure 2.2: Schematic representation of the locations of the different general primer sets (MY 09/11, PGMY09/11, GP5+/6+, SPF1/2 (SPF₁₀) and SPF1/GP6+) on the HPV genome. The circular HPV DNA genome is represented by a single line, and the boxes show the positions of the early (E) and late (L) genes. The position of the different primer sets is indicated within the conserved L1 region. The amplicon sizes are also indicated. PGMY09/11 and MY09/11 have the same amplicon size, 450bp. The SPF1/GP6+ primers were modified from the GP5+/6+ and the SPF1/2 primer sets. SPF1/2 is also referred to as the SPF10 system (modified from Kleter *et al.*, 1999; Huang *et al.*, 2006).

2.2.2.1 MY09/11 and PGMY09/11 primers

The MY09/11 primers amplify 27 HPV types, which were initially detected by the HPV dot blot (Manos *et al.*, 1989). HPV detection using the reverse line blot hybridisation was developed to improve simultaneous detection of multiple HPV types (Gravitt *et al.*, 1998). This detection method is detailed in section 2.3.3. The efficiency of amplification using the MY09/11 primer set varies among the different HPV genotypes due to the degeneracy of the primers. Improved amplification of HPV types was achieved by the design of the PGMY09/11 set. These eliminated degeneracies found in the MY09/11 set, thereby improving the sensitivity, specificity and reproducibility of the PCR (Gravitt *et al.*, 2000; Coutlée *et al.*, 2002). PGMY09/11 primers amplify 37 HPV types compared to the original 27 types amplified by the MY09/11 set. The commercially available Roche Linear Array HPV Genotyping Test (Roche Diagnostics) incorporated the PGMY09/11 primers for HPV amplification.

2.2.2.2 GP5+/6+ primers

The GP5+/6+ primers were designed as a modification to the GP5/6 primers (de Roda Husman *et al.*, 1995). The GP5/6 primers were elongated by three and five nucleotide bases, respectively, at the 3' terminals to result in the GP5+/6+ primers. The post-amplification detection method is Southern hybridisation of amplicons to known HPV type-specific probes. These primers detect 22 HPV types with improved sensitivity compared to the GP5/6 primers (de Roda Husman *et al.*, 1995). Development of HPV detection assays led to the use of the GP5+/6+ PCR enzyme immunoassay (EIA) (section 2.2.2.2.1) and the reverse line blot hybridization (Jacobs *et al.*, 1997; van den Brule *et al.*, 2002). Recent improvements on the GP5+/6+ primers resulted in the Broad-Spectrum GP5+/6+ (BSGP5+/6+) primers, which target 27 HPV types (Schmitt *et al.*, 2008). The primers were designed by alignment of 48 HPV genotypes. When compared to the GP5+/6+ primers, the BSGP5+/6+ set shows higher detection rates for some HPV types, for example, HPV 30, 39, 42, 44, 51-53, 68, 73 and 82 (Schmitt *et al.*, 2008).

2.2.2.2.1 Enzyme immunoassay

After the PCR with the GP5+/6+ primers, the amplicons are detected by hybridisation to digoxigenin (DIG)-labeled HPV-specific probes. This assay can also be used with amplicons generated by the SPF1/2 primers. The probes target a broad spectrum of high- and low-risk HPV types and for each primer set the probe mixture solutions are different. The double-stranded amplicons are captured onto streptavidin-coated microtiter plate in the presence of hybridisation buffer. Denaturation of the captured amplicons is achieved under alkaline conditions using NaOH (Kleter *et al.*, 1998). DIG-labeled HPV probes in hybridisation buffer are added to the microtiter plates containing the captured denatured amplicons. This allows hybridisation of the relevant HPV sequences to the specific HPV probes. The bound amplicons are detected by use of alkaline phosphatase-conjugated anti-DIG antibody and alkaline phosphatase substrate (p-Nitrophenyl phosphate) (Kleter *et al.*, 1998; Söderlund-Strand *et al.*, 2005). The reaction between alkaline phosphatase and its substrate is stopped with acid and the optical density (OD) readings are determined. The specimen is considered positive when the OD values are 2.5 times above the PCR negative control (Kleter *et al.*, 1998). This assay detects the presence of HPV DNA from a total of 38 types without distinguishing the genotypes when used with

GP5+/6+ primer generated amplicons (Söderlund-Strand *et al.*, 2005; Hesselink *et al.*, 2008).

2.2.2.3 SPF1/2 (SPF10)

These primers were designed in a similar target region of the L1 as the MY11 and GP5+ primers but targeting a shorter amplicon of 65bp (Kleter *et al.*, 1998). The post-amplification detection methods are Southern hybridisation, EIA and the line probe assay, INNO-LiPA HPV prototype research assay (LiPA) (Kleter *et al.*, 1998; Kleter *et al.*, 1999). The EIA on amplicons generated by SPF1/2 primers detects a minimum of 74 types without distinguishing the HPV types. The LiPA detection system is similar to the reverse line blot hybridisation method, except that the primers used for HPV amplification prior to detection are the SPF1/2 set (Kleter *et al.*, 1999). The mixture of probes is also different to those in reverse line blot hybridisation: LiPA detects and distinguishes 25 HPV types. Among the 25 types identified by the LiPA, only three HPV types (30, 57 and 74) are not detected by the reverse line hybridisation after amplification with PGMY09/11 primers. The LiPA can also be applied to amplicons generated by the MY09/11 primers (Kleter *et al.*, 1999).

The SPF10 primers are reported to have increased sensitivity for HPV DNA compared to the GP5+/6+ set. As a result, they are capable of detecting HPV in specimens with low viral load (Kleter *et al.*, 1998). The increased sensitivity may be due to the short amplicon compared to the other PCR systems, since the kinetics of PCR generally favour short amplification products (Kleter *et al.*, 1999). When considering the detection of HPV in multiple infections, the PGMY09/11 and the SPF multiple primers were shown to be more robust for HPV detection than the consensus MY09/11 primers (Kleter *et al.*, 1998).

2.2.2.4 Type-specific primers

In some cases, HPV type-specific primers have been used for the detection of HPV DNA. The primers are designed to amplify sequences within the E6/E7 or L1 regions (Lin *et al.*, 2008). The amplicons are subjected to agarose gel analysis and sequencing for HPV detection or detected using the EIA (Castle *et al.*, 2008). The main advantage of type-specific PCR is the higher sensitivity for HPV DNA when compared to the other forms of PCR. However, the use of type-specific PCR assays alone in epidemiological studies is

impractical due to the diversity of anogenital HPV (Lin *et al.*, 2008). Therefore, the major disadvantage of this method in diagnostics is the need for parallel multiple PCR reactions to detect multiple HPV types in a single sample (Lin *et al.*, 2008). The process is labour intensive, especially when working with many samples as is the case in epidemiological studies. This method is most useful when used as an additional method, providing concrete and further evidence on the presence of particular HPV types.

2.2.3 HPV distribution in South Africa

When classified according to cervical disease, HPVs 52, 53, 16 and 35 were more common in 57 Cape Town women with LSIL while 53 women with HSIL were more commonly infected with types 16, 35 and 18 (Allan *et al.*, 2008). In contrast to these results, a study involving eight South African women with HSIL from Pretoria showed that the commonly prevalent types were HPV 35, 58 and 66 (Said *et al.*, 2009). In both these studies the Roche Linear Array HPV Genotyping Test was used for HPV detection. Therefore, the differences in HPV distribution pattern suggest that HPV distribution depends on lesion types and geographical location. A different pattern of HPV distribution was observed among women infected with HIV. In a study in Cape Town involving 37 women with HIV co-infection, HPVs 16, 45 and 66 were reported as the most common types among the HIV positive women (Marais *et al.*, 2008) (section 1.6.3.1). When comparing with the results from 148 HIV positive women from Johannesburg, HPV 16 remained the most common type but HPV 45 was not one of the common genotypes (Firnhaber *et al.*, 2009). In addition to HPV 16, HPV 35, 51 and 66 were also commonly detected in the Johannesburg population (Firnhaber *et al.*, 2009). Based on the differences in distribution among HIV positive and HIV negative women in South Africa, it is clear that HPV distribution pattern is influenced by HIV co-infection, geographic location and lesion types.

In this chapter we evaluated the distribution of HPV types in 109 women from Cape Town who were also infected with HIV. The detection of HPV types infecting these women is crucial in the successful management of women infected with both HIV and HPV in South Africa since HIV infected women are reported to have an increased spectrum of HPV infection. It is also important for the development of prophylactic vaccines that include the most prevalent HPV types in HIV positive women.

2.3 Materials and Methods

2.3.1 Study population

The participants were recruited from an Anti-Retroviral (ARV) treatment clinic in Cape Town, South Africa. All the women in the study were eighteen years or older and the median age was 31 years (range 20-60 years). Informed consent was obtained from interested individuals to whom the details of the study were explained by the research nurse. Those who declined to participate were excluded from the study. Pregnant women, those who had had hysterectomies or were menstruating at the time of recruitment, were excluded from the study. A total of 109, both black and mixed race women, were involved in the study. All of the women had a Pap smear and those with HSIL were referred for colposcopy while the necessary follow-up was arranged for those with LSIL and ASCUS (Moodley *et al.*, 2009). Dr Jennifer Moodley was responsible for the recruitment of participants, Pap smear tests and also arranged for the necessary follow-up of women in this study. The study was approved by the Human Ethics Committee at UCT in 2005 and the ethics approval number is 068/2005.

2.3.2 DNA extraction

Frozen cervical scrapings stored in Digene transport medium (Qiagen, Gaithersburg, USA) were thawed and DNA extraction performed using the MagNA Pure Compact Nucleic Acid Isolation Kit I (Roche Diagnostics, Mannheim, Germany). An automated extraction procedure was performed using a MagNA Pure Compact Instrument (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's instructions. Reagent cartridges, tip trays, elution tubes were placed into respective positions. The cervical samples (400µl) in 2ml tubes were loaded onto the sample tube rack. Sealed reagent cartridges containing the following reagents were used: Proteinase K, lysis buffer, magnetic glass particles (MGPs), isopropanol, wash buffer I and II, and elution buffer. Lysis buffer containing guanidine salts was used to lyse the cells. Proteinase K and the salts were also used to denature contaminating proteins, including nucleases. The DNA was then immobilized onto the MGPs and precipitated by the addition of isopropanol. Denatured proteins and other contaminants were removed by washing with buffers containing absolute ethanol. DNA was released from the beads at 80°C and eluted into 2ml sterile tubes in 200 µl of elution buffer (nuclease-free, sterilised water). Extracted DNA samples were stored at -20°C until further use. The use of nuclease-free sterile tips

and tubes eliminated contamination of DNA. The extraction was performed in an area dedicated for nucleic acid extraction only.

2.3.3 PCR amplification and HPV genotyping

The Roche Linear Array HPV Genotyping Test (Roche Diagnostics, Mannheim, Germany) was used for the detection of HPV types in specimens. This test was used in this work because of its ability to distinguish specific HPV types within the high- and low risk groups, which is not possible with the HC2 test. Preparation of specimens for PCR was performed in a PCR-clean area to avoid contamination with amplified products. Amplification and detection of HPV types were performed in the post-amplification area. DNA from clinical specimens was amplified according to the instructions of the HPV Genotyping and Detection Kit (Roche Diagnostics, Mannheim, Germany). The consensus primer set (PGMY09/11) designed in the conserved L1 region (Figure 2.2) was used to amplify a broad spectrum of 37 HPV genotypes in a single PCR. The reverse primers were biotin labeled at the 5' end to enable capture of the alkaline denatured amplicons onto streptavidin coated strips during post-amplification detection. The HPV master mix contained the PGMY09/11 primers to amplify HPV DNA from 37 types (450bp) and GH20/PC04 primers to amplify 268bp of the β -globulin gene: the human β -globulin gene was co-amplified in the same PCR reaction to assess sample adequacy, DNA extraction and amplification efficiency. The master mix also contained the enzymes AmpliTaq® Gold DNA polymerase and uracil-N-glycosylase (AmpErase); dNTPs including dUTPs. Master mix reagents in a total volume of 580 μ l were mixed with 125 μ l of Magnesium chloride solution to give a working HPV master mix. The Mg^{2+} in solution was for enhanced activity of AmpliTaq® Gold DNA polymerase. This enzyme extended the primers bound to target DNA for the amplification of HPV and human DNA. The AmpErase enzyme eliminated carry-over contamination in PCR by catalysing destruction of amplicons containing dUTP. This ensured selective amplification of the required target DNA only. Fifty microlitres of DNA sample were added to 50 μ l aliquots of working HPV master mix and amplified in the amplification and detection area.

PCR was performed under the following conditions: 50°C for 2 min to activate AmpErase, 95°C for 9 min, followed by 40 cycles consisting of 95°C for 30 sec, 55°C for 1 min, 72°C for 1 min and a final hold at 72°C for 5 min on an AB9700 machine (Applied Biosystems). The reaction was kept at 72°C until the addition of an alkaline denaturation

solution from Linear Array Detection Kit (Roche Diagnostics, Mannheim, Germany) to denature the amplicons and AmpErase. Positive and negative controls from the kit were included in each amplification procedure. The positive control (50µl) had HPV 16 DNA and negative control (50µl) had no HPV DNA. The post-amplification detection of 37 HPV types (HPV6, 11, 16, 18, 26, 31, 33/35/52/58, 39, 40, 42, 45, 51, 53-56, 59, 61, 62, 64, 66-73, 81-84, IS39 and CP6108) was carried out using Linear Array Detection Kit (Roche Diagnostics, Mannheim, Germany). IS39 is a subtype of HPV 82 and CP6108 is also known as HPV 89 (Clifford *et al.*, 2005). The denatured amplicon mixture (75µl) was transferred to appropriate wells in a typing tray containing hybridisation buffer and a single Linear Array HPV Genotyping Strip included in kit. The strip was coated with HPV and β -globulin oligonucleotide probes immobilized onto membrane. Each amplicon was hybridised to the strip at 53°C and was bound complementary probe with a matching sequence. There was one cross-reactive oligonucleotide probe that hybridised with HPV 33, 35, 52 and 58. The Roche Linear Array HPV Genotyping test does not directly detect HPV 52, instead it combines a set of probes detecting HPV 33, 35, 52 and 58 as a group. The specimens that tested negative for HPV33, 35 and 58 individually but tested positive for the group were considered to be HPV 52 positive. Specimens that tested positive for the group and for HPV33, 35 and/or 58 had uncertain HPV 52 status although they were considered to be HPV 52 negative in this study. Following hybridisation and stringent washes with buffers containing SDS and sodium salts, streptavidin-horseradish peroxidase conjugate (SHPC) was added to the strip to bind to the biotin-labeled amplicons hybridised to the probes on strip. A substrate mixture of hydrogen peroxide (H₂O₂) and 3,3',5,5'-tetramethylbenzine (TMB) was added to each strip. The TMB substrate was oxidised to a blue colour by catalysis action of streptavidin-horseradish peroxidase in the presence of H₂O₂. The blue colour precipitated at probe positions and allowed for the reading of strips.

2.4 Results

HPV genotyping was performed in order to evaluate the distribution of HPV types in 109 South African women infected with HIV. The high-risk types tested for were HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 59, 68, 73, 82, IS39 – a subtype of 82 and the probable high-risk HPV 26, 53 and 66. For analysis, the probable high-risk types were grouped under high-risk. The low-risk HPV 6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81 and 89 were

tested for. Other types of undetermined risk were also tested for; HPV 62, 64, 67, 69, 71, 83 and 84 (Section 1.5.1, Table 1.3). For the analysis of results, the seven 'unclassified' types were grouped with the low-risk types. The HPV types were shown as blue bands on HPV genotyping strips (Figure 2.2).

From a population of 109 HIV positive women, 11% (12/109) were found to be negative for HPV DNA (Table 2.1). A total of 37 genotypes were tested for, using the Roche Linear Array HPV Genotyping test and 95% (35/37) of the HPV types were detected in this population. All of the 18 high-risk HPV types mentioned above were present. Table 2.1 shows the distribution of HPV genotypes in this population. Among the 89% (97/109) who were HPV positive, 22% (21/97) had a single HPV type and 78% (76/97) had at least two HPV types. In this study, 14% (15/109) and 15% (16/109) of the women had two and three HPV types, respectively (Table 2.2). Multiple HPV infections were common, with a median number of three HPV types per woman. The range of HPV types was 0-12. The most common HPV types among all women, in order of decreasing frequency, were HPV 61, 66, 53, 58, 18, 45, 70, 35, 16 and 51 (Table 2.1).

A total of 98/109 women had a Pap smear result and 66% (65/98) of them had an abnormal Pap smear result. Among these women the most common high-risk HPV types were 16, 58, 53, 51 (all at 29%), 66 (25%), 18 (22%) and 45 (20%). Among the women with abnormal Pap smear, 10/98 (10%) had HSIL. The high-risk HPV types most prevalent in HSIL were HPV 45 (30%), 16, 35, 39, 58 and 51 all at 30%, as well as types 66, 18, 31 and 33 with 20% prevalence rate (Moodley *et al.*, 2009).

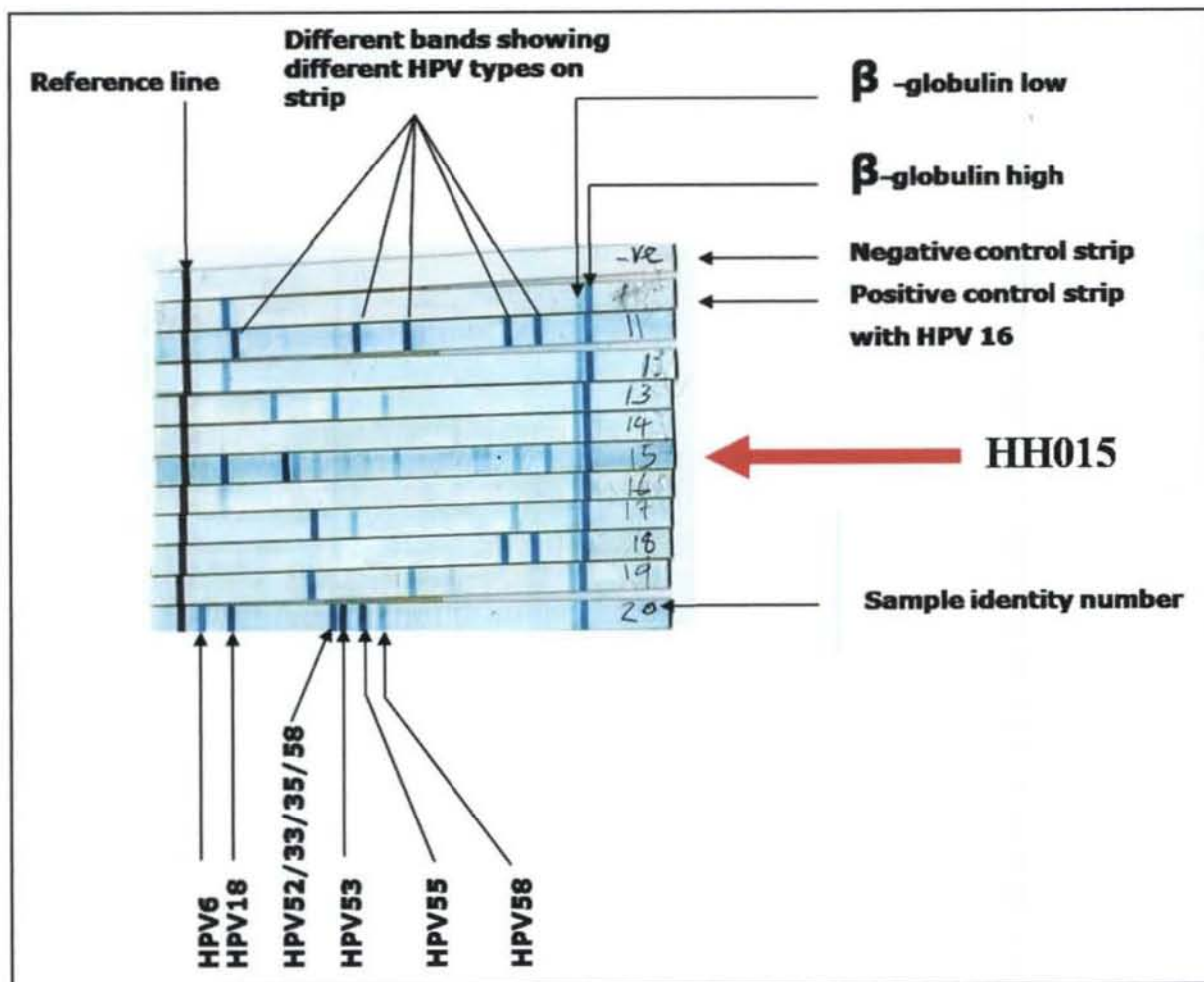







Figure 2.2: Roche Linear Array HPV genotyping test results for 10 clinical specimens, each strip represents a specimen except for the negative and positive control strips. All samples showed both the β -globulin high and low for the results to be validated. The HPV types in each specimen are shown by the blue bands. As an example in this set of results, specimen HH015 (indicated by the red arrow) had, from left to right, HPV 16, 39, 40, 45, 52, 53, 55, 59, 70, 71, 81, and 84. The intensity of the band is directly proportional to the viral load in specimen. For specimen HH015, bands for HPV 16 and 39 were more intense than the others, therefore it was assumed that there were higher viral loads for these two than the other types. The HPV types for specimen HH020 are indicated as HPV 6, 18, 53, 55 and 58. HPV 52 presence in this sample could not be ascertained; therefore the appearance of a band on the group types was inferred to indicate presence of HPV 58.

Table 2.1: Genotyping results of HIV positive women from a clinic in Cape Town, South Africa.

Description	Number of women (n= 109)	Percentage %
HPV negative	12	11.0
HPV positive	97	89.0
High-risk types ^a	89	82.0
66	20	18.3
58	19	17.4
53	18	16.5
45	17	15.6
18	17	15.6
35	16	14.7
16	15	13.8
51	14	12.8
68	12	11.0
39	11	10.1
52	10	9.2
73	10	9.2
33	9	8.3
59	8	7.3
56	6	5.5
31	6	5.5
IS39 ^c	6	5.5
82	6	5.5
26	5	4.6
Low-risk types ^b	8	7.3
61	26	23.9
70	17	15.6
62	13	11.9
54	12	11.0
55	12	11.0
83	11	10.1
84	11	10.1
C P6108 ^c	11	10.1
81	10	9.2
71	7	6.4
72	6	5.5
6	5	4.6
40	5	4.6
69	4	3.7
11	3	2.8
42	2	1.8
67	1	0.9

Key to five most common types

	HPV 61
	HPV 66
	HPV 58
	HPV 53
	HPV 70

^a Due to multiple infections, some individuals had both high- and low-risk types therefore number represents individuals with either high-risk types only or high- and low-risk HPV types. The probable high-risk types were included in the high-risk category. ^b Women with only the low-risk HPV types. ^c IS39 is a subtype of HPV 82 (Clifford *et al.*, 2005). Both IS39 and HPV 82 were detected in one specimen HPV CP6108 is also known as HPV cand89 (Clifford *et al.*, 2005). 'cand' refers to HPV type identified by cloning and characterisation of PCR amplicons (de Villiers *et al.*, 2004). For further analysis, IS39 and HPV 82 were recorded as HPV 82 and CP6108 was recorded as HPV 89.

2.4.1 Low-risk HPV types

The low-risk HPV 6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81 and 89 (CP6108) were tested for. The other types tested for were of undetermined risk and these were HPV 62, 64, 67, 69, 71, 83 and 84 (Section 1.5.1, Table 1.3). These two groups were considered as low-risk types in this study. HPV 61 occurred most frequently in this population, with a prevalence rate of 24% (26/109). The other low-risk types detected in the study, in order of decreasing frequency, were HPV 70, 62, 54, 55, 83, 84, 89, 81, 71, 72, 40, 6, 69, 11, 42, and 67, as illustrated in Figure 2.3. Of all the low-risk HPV types tested for, 94% (17/18) of them were detected. No samples were positive for HPV 64.

2.4.2 High-risk HPV types

The following HPV types in the Roche Linear Array HPV genotyping test were considered to be high-risk: 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73 and 82 (section 1.5.1, Table 1.3). HPV 66 is classified as a probable high-risk type therefore it was grouped together with the high-risk types, as were HPV 26 and 53. All of these types were detected in these women as shown in Figure 2.3. HPV 66 accounted for the highest prevalence among the high-risk HPV types (18%). The common HPV types in the high-risk category were 66 (18%), 53 and 58 (17%), 18 and 45 (16%), 35 (15%) and 16 (14%) (Table 2.1).

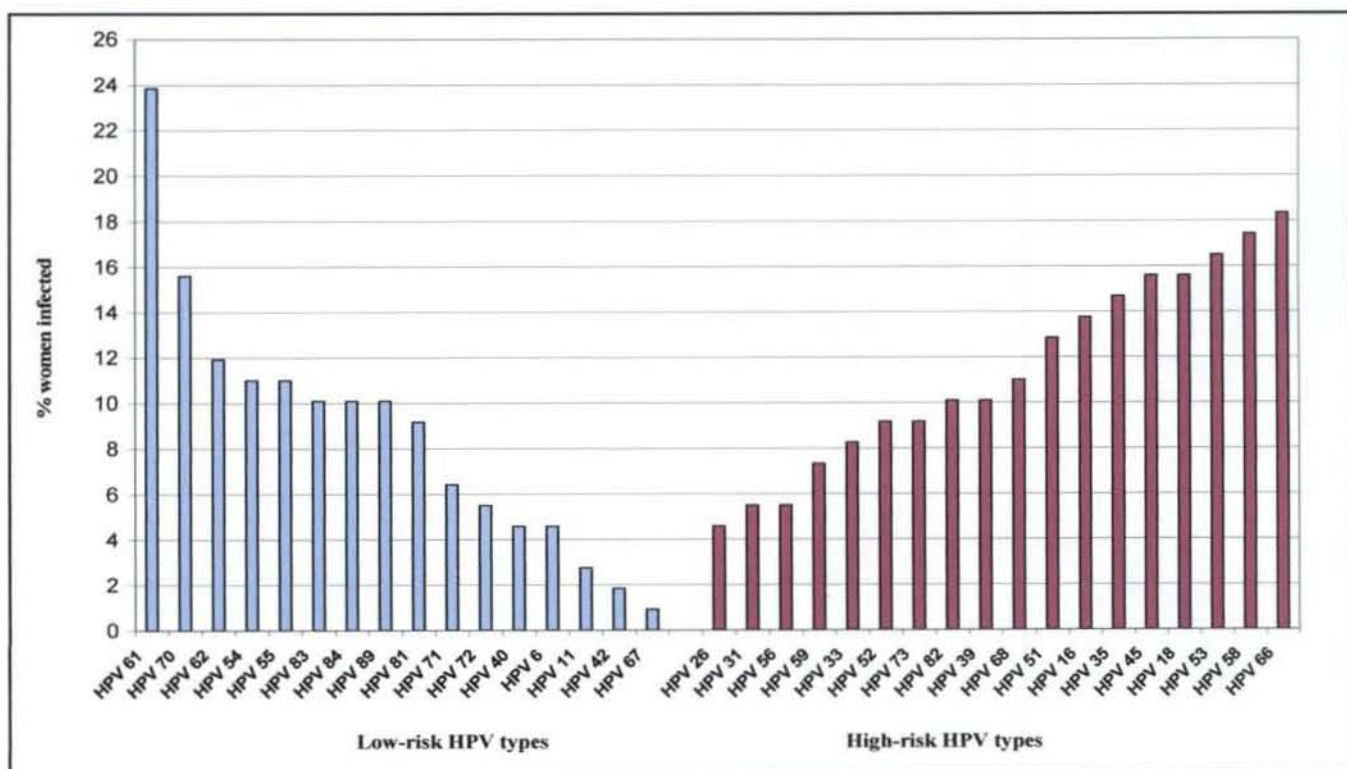


Figure 2.3: The prevalence of low- and high-risk HPV types among 109 HIV positive women at a clinic in Cape Town. The probable high-risk types 26, 53 and 66 were included among the high-risk HPV types. The total percentages were more than 100 because some women were counted more than once due to multiple infections.

Table 2.2: Summary of the prevalence of HPV and multiple infection in HIV positive women (n=109) in Cape Town.

Number of HPV types	Specimens	
	N	%
0	12	11
1	21	19
2	15	14
3	16	15
4	11	10
5	8	7
6	11	10
7	2	2
8	6	6
9	3	3
10	1	1
11	2	2
12	1	1

2.5 Discussion

From this study, we determined that 89% (97/109) of the women were infected with one or more HPV types (Table 2.1). The method used for detection of HPV DNA has a major impact on the perceived diversity of HPV identified in a population. The use of PGMY09/11 primers and reverse line blot hybridisation method has been shown to be an effective way to detect HPV DNA in cervical specimens (Gravitt *et al.*, 2000).

The prevalence pattern observed in this group is similar to that observed in a study in Uganda (Banura *et al.*, 2008) on HIV positive women. Both studies confirm that there is a high rate of HPV infection among HIV positive women. The Ugandan study showed that HIV positive women have higher HPV infection rates as well as increased multiple infections than the HIV negative women. The number of HIV positive women in that study was 82/950 and 88% (72/82) of HIV positive women were HPV DNA positive compared to 73% (635/868) among the HIV negative women (Banura *et al.*, 2008). Levi and colleagues in Brazil also observed a high HPV prevalence of 98% (204/208) among HIV positive women (Levi *et al.*, 2002). Similarly, the HPV prevalence in our study was 89% (97/109), but HIV negative women were not included in our study. Multiple infections accounted for 78% (76/97) of all HPV positive women compared to only 22% (21/97) women with single infections. In the Ugandan study, 65% (53/82) of the HIV positive women had multiple infections, which was an increase in multiple infections from 37% (19/868) among the HIV negative women (Banura *et al.*, 2008). This clearly indicates that HIV positive women have more multiple infections than the HIV negative women. Most women in our study (82%, 89/109) were infected with at least one high-risk HPV type and 7% (8/109) of women had only low-risk HPV types (Table 2.1). When considering that high-risk HPV types can lead to cervical disease or neoplasia, co-infection with high-risk types in HIV positive women might explain why HIV positive women are more susceptible to developing cervical disease. The high rate of infection with high-risk HPV types noted in our study increases the risk of developing cervical neoplasia among HIV positive women.

HIV infected women are more likely to shed HPV over a long period of time than HIV negative women (Baseman and Koutsky, 2005). This implies that HPV detection in these

women is also higher than in the HIV negative women (Banura *et al.*, 2008; Singh *et al.*, 2009). In a study of 208 HIV positive women in Brazil, 98% (204) of them were positive for one or more HPV types (Levi *et al.*, 2002). This may suggest that HIV positive women would have at least one HPV type detected. In our study, 12% (11/109) of the women had no HPV DNA detected. While this result may be a true reflection of HPV status in these women, it is also possible that HPV DNA detection was hindered by HPV DNA integration. Integration of HPV DNA in HSIL and cervical cancer can disrupt the primer binding sequences within L1 ORF or can lead to loss of the L1 ORF, therefore leading to a negative result by PCR (Clavel *et al.*, 1998; Walboomers *et al.*, 1999).

A negative HPV DNA result can also be caused by latency. Latent HPV DNA can exist at very low copy numbers that are difficult to detect even with sensitive methods like PCR. Latency of HPV DNA is not fully understood. Therefore, there is still need for further investigation into how long latency lasts, and when or under what conditions is HPV DNA re-activated to detectable levels (Schiffman and Kjaer, 2003). HPV latency could result in reactivation of latent HPV DNA in HIV infected individuals, thus explaining the wide range of HPV types detected.

In contrast to other epidemiological studies in a worldwide general population, HPV 16 was not the most common HPV type identified among HIV positive women (with or without cervical abnormalities) in our study. HPV types other than 16 have been detected commonly among women with HIV infection. This HPV distribution pattern was observed in studies in Brazil, America, Uganda and Italy (Levi *et al.*, 2002; Luque *et al.*, 2006; Banura *et al.*, 2008; Tanzi *et al.*, 2009). Among all the women in our study, HPV 16 prevalence was only 14% (15/109). HPV 61 (24%, 26/109), a low-risk type, was found to be the most prevalent type among women with normal and abnormal cytology (Table 2.1). Tanzi and co-workers (2009) also reported in their study on HIV infected individuals in Italy that HPV 61, together with 6 and 11, were common HPV types among all HIV positive women in that population (Tanzi *et al.*, 2009). This suggested that low-risk types may persist in HIV positive women. While this may seem extraordinary in HIV negative individuals, it is not surprising that low-risk infections tend to persist in HIV positive women. This is due to the low HPV clearance in HIV infected women as a result of reduced HPV-specific immune response (Palefsky, 2006). The high-risk HPV types 56, 52, 53 and 58 were frequently detected in women with HIV

infection (Luque *et al.*, 2006; Banura *et al.*, 2008). In agreement with results from these and other studies in South Africa, our study also observed an increase in other high-risk HPV types other than HPV 16. Overall, HPV 66, 58, 53, 45, 18 and 35, in order of decreasing frequency, were frequently observed in our study. All of these HPV types have been reported in similar studies in South Africa (Marais *et al.*, 2008, Firnhaber *et al.*, 2009), although the order of frequency was different across studies.

Due to immunosuppression, HIV positive individuals are susceptible to frequent and persistent HPV infections and the immune system may not be able to clear low-risk HPV infections (Chirenje, 2005; Frisch *et al.*, 2000). As a result, HIV positive individuals experience an increased rate of progression to CIN2/3 and the low-risk types may also lead to high grade lesions in severely immunosuppressed women (Luque *et al.*, 1999; Frisch *et al.*, 2000; Clifford *et al.*, 2006b). This was observed from a meta-analysis on HPV infection among HIV positive women from different countries, in which the low-risk HPV types 11 and 61 were more frequently detected in some of the HIV infected women with HSIL. No high-risk HPV types were detected in 9% (16/185) of the women with HIV infection who had HSIL (Clifford *et al.*, 2006b). Therefore, this gives possibility that low-risk HPV types can cause cervical disease in immunocompromised women.

When considering cervical disease, the most common HPV types in women with abnormal cytology in this study were HPV 16, 58, 53, 51, 66, 18 and 45 (Moodley *et al.*, 2009). Among these women the most common high-risk HPV types were 16, 58, 53, 51 (all at 29%), 66 (25%), 18 (22%) and 45 (20%). The prevalence of HPV 16 among these women did not dominate over the other types as the prevalence was the same as for some other high-risk types (Moodley *et al.*, 2009). This is in agreement with the general trend that HPV 16 is not predominantly prevalent in HIV infected women. These results were comparable to another study in Cape Town where HPV 45, 66 and 16 were also found to be common in HIV infected women (Marais *et al.*, 2008). In yet other studies in Uganda and Rwanda, HPV 52, 58, 51 and 16 were common in HIV infected women with disease (Blossom *et al.*, 2007; Singh *et al.*, 2009). Among these, HPV 52 was the only type that was not reported among the common HPV types in our study.

From the results in different countries, it is clear that the spectrum and prevalence of HPV is different in HIV positive women from different geographic locations. However, differences in the prevalence rates can also be due to the sample sizes and methods used in genotyping. For HPV genotyping, Singh and co-workers (2009) used the SPF10 primers followed by Southern hybridisation and dot blot hybridisation (Singh *et al.*, 2009). Blossom and colleagues (2007) used the PGMY09/11 primers followed by reverse line blot hybridisation for detection of 27 HPV types (Blossom *et al.*, 2007). In our study, HPV genotyping was performed with the Roche Linear Array HPV Genotyping test (section 2.3.2), which does not distinguish HPV 52 from HPV 33, 35 and 58. The population sizes of HIV infected women in Ugandan and Cape Town studies were small (37/106 and 37/115, respectively) compared to our study of 109 HIV positive women. Singh and colleagues' study had 710/936 HIV positive women, a larger population compared to our study in 109 women. Therefore the differences in the genotyping methods and sample sizes could have contributed to the differences in HPV 52 patterns in the different studies. On the other hand, it is also entirely possible that there are different genotypes in the populations under study.

The prevalence of HPV 52 (9%) in our study may not give a true reflection of its prevalence among HIV women as it was grouped with HPV types 33, 35 and 58. Because HPV 52 was scored positive when neither of these specific types was present individually, co-infection with HPV 52 was possible for each specimen that was positive for the group as well the other types. As a result, the actual prevalence of HPV 52 could be higher than the observed 9% (10/109). This underestimates the existence of this type in a population. A further analysis of the specimens presumed to be HPV 52 positive by use of type 52 specific primers could improve the detection of this genotype and therefore give more reliable information on its prevalence. The use of one detection method for HPV detection in prevalence studies may result in the underestimation of the prevalence of some HPV types in the population as opposed to two or more methods.

This study showed that there is a wide range of HPV types infecting HIV positive women. A total of 35 low- and high-risk HPV types were identified. There were frequent multiple infections detected, with 70% (76/109) of the women showing at least two HPV types. The median HPV types per woman was three, ranging from 0-12 types.

Chapter 3

Analysis of HPV 16 variants

3	INTRODUCTION.....	50
3.1	HPV VARIANTS	50
3.2	MATERIALS AND METHODS.....	53
3.2.1	<i>PCR amplification of HPV 16 LCR.....</i>	<i>53</i>
3.2.2	<i>Cloning and sequencing.....</i>	<i>54</i>
3.3	RESULTS.....	55
3.3.1	<i>PCR amplification of HPV 16 LCR.....</i>	<i>55</i>
3.3.2	<i>Sequence analysis of LCR.....</i>	<i>55</i>
3.3.3	<i>Phylogenetic classification of HPV 16 variants.....</i>	<i>58</i>
3.4	DISCUSSION.....	61

3 Introduction

3.1 HPV variants

HPVs are divided into genera, species, types, subtypes and variants. Members of different genera (alpha-, beta- and gamma-papillomavirus) share less than 60% similarity in nucleotide sequence in the capsid protein L1 ORF. Same species groups share between 60-70% similarity in the sequence of the same gene. HPV types within a species are defined by 71-89% nucleotide sequence similarity within the L1 ORF (IARC, 2007). Subtypes share between 90-98% homology with prototype sequence (Hazard *et al.*, 2007). This means that subtypes differ from the prototype sequence by 2-10% within the L1 ORF. Variants show 98-99% homology in DNA sequence within the same region, which is less than 2% difference from the prototype in the nucleotide coding region of L1 protein (de Villiers *et al.*, 2004). The difference in the L1 ORF of prototype and variant sequences can be in one or more nucleotide bases. These point mutations may either be silent or result in a change in an amino acid when compared to the protein sequence of the prototype HPV sequence (Tu *et al.*, 2006).

Nucleic acid sequencing data from previous studies show that there are a number of naturally occurring 'variants' of certain HPV types. Most studies have been done on the intratypic variants of HPV 16 since it is one of the most commonly occurring high-risk HPVs. HPV 16 variants seem to be more common than variants of other HPV types and are normally associated with particular ethnic groups (Hazard *et al.*, 2007), and geographical regions (Kurvinen *et al.*, 2000). While some authors argue that subtypes appear to be rare or virtually absent compared to variants, Hazard and others report that HPV subtypes are common (Bernard, 2005; Hazard *et al.*, 2007). According to Hazard and colleagues (2007), about 47 subtypes/putative subtypes of HPV exist. These are classified into 30 subtypes within the genus beta-papillomavirus, ten within genus gamma-papillomavirus and seven subtypes within the genus alpha-papillomavirus (Hazard *et al.*, 2007).

The distribution of HPV 16 variants across five continents is grouped according to their location: Asian-American (AA) found in the Central and South America; African (Af) in African continent; Asian (A) predominant in South East Asia and lastly European (E)-

found in all regions except Africa (Berumen *et al.*, 2001). Some variants are more often associated with cervical cancer than others. In fact, Berumen and others (2001) reported that infections with the AA variants had an increased risk of developing cervical cancer compared to those with E variants (Berumen *et al.*, 2001). Tu and colleagues (2006) claim that some Af variants are associated with increased risk of invasive cervical cancer compared to the E variants in South African women (Tu *et al.*, 2006). In contrast to this, Tornesello and colleagues (2000), in their study among Ugandan males, argue that there is no evidence to suggest different oncogenic activity in the Af-1 HPV 16 variants as compared to the European variants. However, they do report that there is enhanced transcription activity in the LCR of Af-1 variants as compared to that of E variants (Tornesello *et al.*, 2000). Interestingly, after comparing work that had been done previously on other variants in USA, Brazil, Costa Rica, China, it was shown that association with cervical cancer depends on the geographical location (Tu *et al.*, 2006). Whether or not a certain variant is associated with increased risk for cervical cancer seems to depend on the possibility that specific populations have developed a certain level of resistance to the transforming potential of variants found commonly in their area (Tu *et al.*, 2006). It is also possible that the absence of immunologic interactions is the major reason for the reduced evolutionary restraints, therefore allowing adaptation of new HPV variants to different types of differentiated epithelia (zur Hausen, 2000).

Genetic variation has been identified by amplification and sequencing of regions within the E6 and E7 genes and the LCR. Nucleotide differences within these regions indicate a variant (Bernard, 2005). HPV 16 variants have been identified in the E7, L1, L2, E5 and E2 genes as well as E6 and LCR. The sequence variations or mutations in these regions have the potential to modify the function of the encoded protein (Berumen *et al.*, 2001) or affect the binding affinity of transcription factors when they occur in the LCR. Mutations in the binding site (s) of negative transcription factor Yin-Yan factor 1 (YY1) were shown to increase expression levels of E6/E7. Furthermore, point mutations in HPV 16 and 18 LCR enhancer regions and upstream of it result in increased transcription activity (Tornesello *et al.*, 2000).

Virus replication and transcription of E6/E7 proteins are regulated by E2 proteins binding to regions within the LCR. The LCR also regulates transcription through a number of *cis*-responsive elements (Tornesello *et al.*, 2000). These response elements include the

steroid responsive elements, particularly the glucocorticoid response element (GRE). The GREs of some HPVs can function as progesterone response elements (PREs), and therefore allow progesterone levels to play a significant role in HPV-induced transformation. It has been demonstrated that several glucocorticoid-regulated promoters within the LCR can be induced by progesterone (Chan *et al.*, 1989). High progesterone levels occur during pregnancy, thus explaining why high parity is considered a risk factor.

The LCR is divided into three distinct sections comprising of the 5', central and 3' segments (Figure 3.1), all of which have different functions. The 5' end has a high A+T content (approx 85%) and contains the negative regulatory element (NRE) as well as the nuclear matrix attachment region (MAR) (Tornesello *et al.*, 2000). The MAR seems to be conserved within most genital HPVs, indicating a significant role in the virus life cycle. The NRE acts at the level of mRNA stability whereas the MAR represses the expression of the E6/E7. The central segment of LCR acts as an epithelial specific transcription enhancer (Tornesello *et al.*, 2000). This enhancer region acts as binding region for cellular transcription factors such as NF-I, NF-IL6, AP-1, YY1, TEF-1 and Oct-1, whose binding has been proven in *in vitro* experiments. The 3' segment houses the origin of replication and the E6/E7 promoter, p97 (Tornesello *et al.*, 2000, Stünkel *et al.*, 2000). The p97 promoter of HPV 16 is regulated through E2 function, binding of transcription factors to the enhancer region and the GRE. While the binding of most of the transcription factors generally stimulates the promoter activity, binding of NF-IL6 and YY1 to the enhancer region can also down-regulate the promoter activity (Kurvinen *et al.*, 2000).

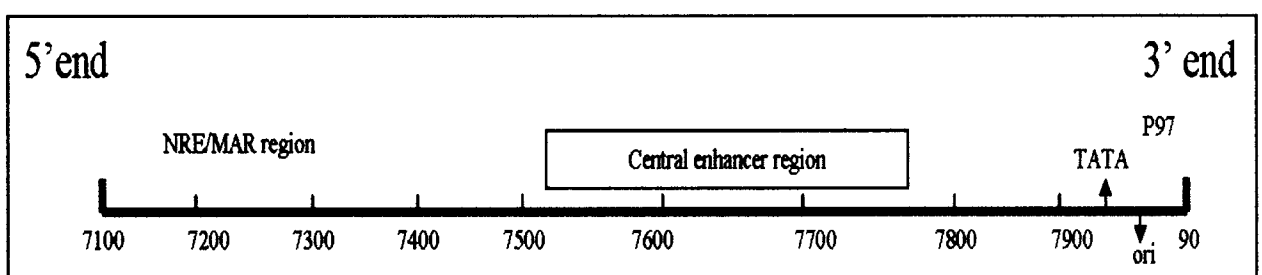


Figure 3.1: Schematic representation of HPV 16 LCR. Transcription factors including Oct-1, NF-1, YY1, and AP1 bind in the enhancer region in addition to other sites along the LCR (Kämmer *et al.*, 2000; Tornesello *et al.*, 2000).

The regulation of viral/cellular genes is critical in determining the outcome of HPV infection as it may lead to cervical cancer. Therefore, any nucleotide changes within HPV enhancer/promoter regions where transcription factors bind may have an impact on disease progression (Tornesello *et al.*, 2000). Sequence changes within the LCR and other HPV genes are associated with the oncogenic potential of the infecting HPV. As an example, the mutation within the YY1 binding site is known to increase transcription of the E6/E7 oncogenes (Pande *et al.*, 2008). Studies on the HPV 16 genome revealed high polymorphism in LCR, resulting in variant identification. In these studies, the prototype HPV 16 genome was identified as an E variant (Kämmer *et al.*, 2002). A good number of studies have been done on HPV 16 variants in other countries, including Finland, Brazil, Mexico and India, but few studies have looked at the HPV 16 variants in African women. Although a lot of work has been done to identify the HPV genotypes in HIV infected women, there is limited information on studies of HPV 16 variants in these women. In a study on HPV genotypes and HPV 16 variants in HIV positive women in Italy, Tornesello and colleagues (2008) identified only the E and Af-2 variants. In that study, HPV 16 variants were identified from HIV positive and HIV negative women with normal cytology and cervical neoplasia (Tornesello *et al.*, 2008).

The primary focus of work reported in this chapter was to identify HPV 16 variants in HIV infected individuals with varying disease severity. From the results in Chapter 2, it was observed that 15/109 (14%) of the individuals were infected with HPV 16. HPV 16 has been identified to have at least five variants which tend to occur according to geographical locations and some of which have been claimed to be more virulent than others. Therefore, HPV 16 variants in our sample set of 15 women were investigated.

3.2 Materials and Methods

3.2.1 PCR amplification of HPV 16 LCR

Fifteen specimens positive for HPV 16 were identified using the Roche Linear Array HPV Genotyping test as described earlier (Chapter 2, Table 2.1). Primers specific to HPV 16 LCR were designed to target the LCR. The primers are shown in Table 3.1. PCR mix was set up in a total volume of 50µl as follows: 20pmol of each primer, 200µM dNTP mixture, 1.3U of high fidelity *Pfu* DNA polymerase enzyme (Promega, Madison, USA),

1X Buffer with magnesium salt (Promega, Madison, USA) to enhance enzyme activity and 3µl DNA. The dNTP mixture contained the nucleotides (dCTP, dTTP, dATP and dGTP) for the extension of primers. The high-fidelity enzyme was used to reduce replication error-rate during the amplification of LCR. The amplification protocol was followed according to Kurvinen *et al* (2000): initial denaturation at 94°C for 10 minutes, 35 cycles consisting of 94°C for 1 minute, 62°C for 1 minute and 72°C for 2 minutes to allow for primer extension. Final extension was at 72°C for 10 minutes for the final elongation of primers.

Table 3.1: Primers used in the amplification of HPV 16 (Kurvinen *et al.*, 2000)

Primer name	Sequence and length of primers	Position	Length (bp)
HPV16LCRF	5'-CCTCATCTACCTCTACAACGCTAAACGC-3' (29)	7108-7136	1028
HPV16LCRR	5'-CGTCGCAGTAACTGTTGCTTGCACTACACAC-3' (31)	222-192	

3.2.2 Cloning and sequencing

The cloning and sequencing of LCR amplicons was performed in order to compare sequence variation with the prototype HPV 16. The amplicons were electrophoresed on 1% agarose gel as detailed in Appendix Aiii. The amplicons were excised from gel and purified as outlined in Appendix Av. The purified PCR amplicons of LCR were cloned using a pGEM®-T Easy cloning system (Promega, Madison, USA) (Appendix Avi) to facilitate sequencing using Sanger's dideoxy chain termination method (Appendix Aviii). The cloned fragments were transformed into *E.coli* DHα cells as outlined in Appendix Avii. Single colonies of clones were grown in LB (Appendix Ai) and DNA was isolated from bacterial cultures according to protocol in Appendix Aii. At least two clones from each specimen were selected for sequencing. Each clone was sequenced twice to eliminate PCR and sequencing artefacts, using the primers in Table 3.1. Sequencing was performed at Macrogen, Korea, using the Sanger chain termination method. The sequences were compared to the reference HPV 16 (GenBank accession number K02718) for the detection of nucleotide variation. Differences in nucleotide sequence were recorded with respect to their position in genome.

3.3 Results

3.3.1 PCR amplification of HPV 16 LCR

The LCR was amplified from HPV 16 positive isolates (n=15) that were identified using the Roche Linear Array HPV Genotyping test. The full LCR was amplified from four samples as shown in Figure 3.2, to give a 1028bp fragment. The same amplicon size (1028bp) was obtained from the other eleven specimens HH015, HH016, HH038, HH049, HH054, HH055, HH065, HH069, HH082, HH095 and HH097. All amplicons were cloned and sequenced for investigation of sequence variation within the LCR.

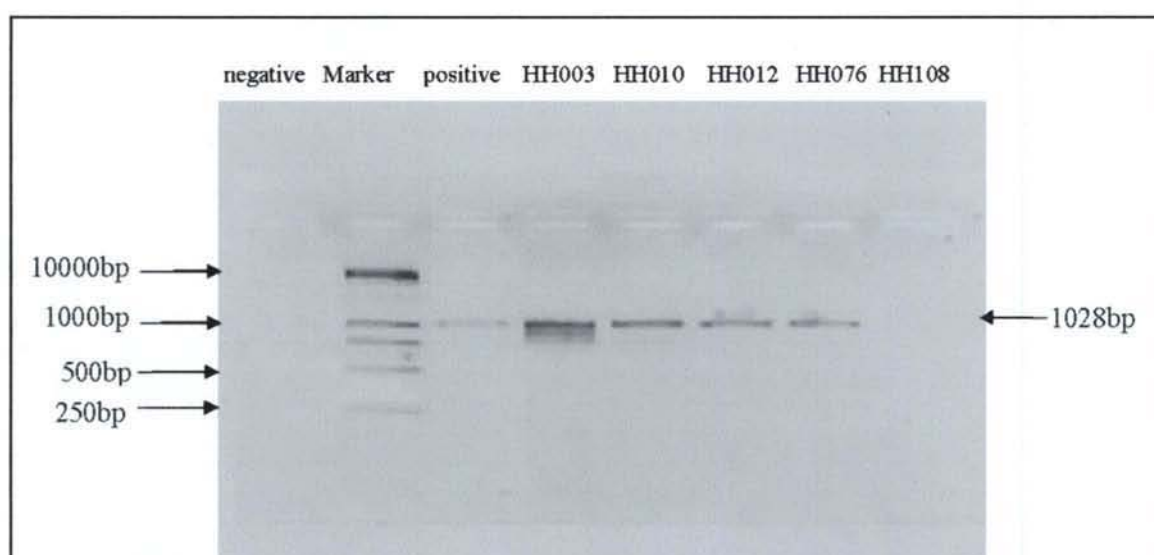


Figure 3.2: Representative agarose gel electrophoresis after amplification of HPV 16 LCR from clinical specimens. This agarose gel shows amplified LCR from four specimens that were HPV 16 positive. Negative lane shows a no DNA negative control and lane HH108 was a DNA negative control. Specimen HH108 was HPV 16 negative. The positive lane was a positive DNA control from specimen HH015.

3.3.2 Sequence analysis of LCR

Sequencing of all cloned amplicons was performed in order to identify basepair changes within the LCR in comparison to the prototype HPV 16. The original HPV 16 sequence published (GenBank accession number K02718) had sequencing errors, which were later corrected. These corrections were within the LCR at nucleotide positions 7432-7433 (CG), which was corrected to CCG and at position 7861 where nucleotide (nt) A was deleted as part of the correction. The corrections resulted in a shift in the numbering of

nucleotide positions by two or three bases. The numbering used in this analysis was according to the original sequence. The prototype sequence used in this study was HPV 16 GenBank accession number K02718.

The nucleotide base differences for the variants identified are shown in Table 3.2. Analysis of HPV 16 variants in these women showed that both European and African variants can exist within the same individual (Table 3.2).

A total of 36 different point mutations were detected within the LCR at 35 different positions. Both the transition and transversion mutations were observed; 44% transition mutations while 56% were transversion point mutations. There were common nt changes noted at 16 positions across most of the specimens. These were at positions G7191T, A7231C, G7385C, G7432A, A7482C, G7486A, G7518A, C7666T, C7686A, C7761T, C7783T, G7823A, G7831T, A7834C, A7836G, and C31T (Table 3.2). Two different transversion mutations at same nt position (A7231C, A7231T) were observed. The number of nt changes in each isolate ranged from 1-17. The most common mutation was a transition point mutation G7518A (highlighted in Table 3.2), which was detected in 73% (11/15) of the specimens. Five nt changes that have not been reported before were also observed. These were A7230G, C7550A, T7865C, C7874A, and G59A. Among these, A7230G and C7874A were detected in the Af-1 variants and the other mutations were observed in Af-2 variants.

Table 3.2: Nucleotide sequence changes within HPV 16 LCR

[illegible]

Single nucleotide base differences within the LCR of HPV 16 variants in comparison to the prototype sequence HPV16R, GenBank accession number K02718. Sample identity refers to the specimen number, and the letters a-j refer to the different cloned sequences per specimen (isolates) analysed per specimen. Total refers to the total number of mutations in each analysed clone. The red highlighted base pair position indicates the position where the most frequent base-pair change among all the isolates. Some of the common binding regions within the LCR are indicated at the respective base positions.

Some of the mutations were observed in transcription factor binding regions of the LCR. G7385C mutation was observed in the GRE region in 9 (60%) isolates while the mutation A7455T, in 3 isolates (20%), occurred in the E2 binding site. The transition mutation C7761T was detected upstream of the NF1 transcription factor binding site while T7711A occurred within its binding site. The mutations G7518A, C7783T and G7823A were found within the YY1 binding site. A transversion mutation, A7633C, which is located at the AP1 binding site, was observed in 2/15 (13%) of the specimens. Other mutations located in binding sites include C7550A, C7666T, C7686T, A7726C, G7796C, C7815T, G7831T, A7834C A7836G and C31T (Tanzi *et al.*, 2009; Pande *et al.*, 2008).

3.3.3 Phylogenetic classification of HPV 16 variants

A phylogenetic analysis of the cloned sequences was performed to identify the variants present in each specimen. The sequences were compared to those of the known variants of HPV 16. The phylogenetic classification of the variants found in the specimens is shown in Figure 3.3. The E variants were present in 7/15 (47%) of the specimens listed in Table 3.3. In addition to the E variants, the A, AA, Af-1 and Af-2 type variants were also identified. Some specimens had more than one variant and multiple infections were noted in 4/15 (20%) of the specimens (HH015, HH054, HH069 and HH082). In two of these specimens (HH054 and HH082), the E variant was present together with Af-2 variant (Figure 3.3). In the third specimen (HH069), AA and Af-2 variants were detected and one specimen (HH015) was co-infected with A and E variants as shown in Figure 3.3.

Table 3.3: The HPV 16 variants identified among the specimens according to phylogenetic distribution.

Variant type	Specimen(s)	Frequency (%) of variants (n=15)
AA	HH069a	7 (1/15)
Af-1	HH012	7 (1/15)
Af-2	HH003, HH016, HH049*, HH054*a, HH055, HH069c, HH076, HH082a, HH095	60 (9/15)
E	HH010*, HH015c, HH038, HH054b, HH065, HH082d, HH097	47 (7/15)
A	HH015d	7 (1/15)

AA- Asian American, Af-1 – African type 1, Af-2 – African type 2, E – European, A – Asian. Letters a-d, after specimen number indicate different cloned sequences per specimen (isolates) within the same specimen or sample as shown in Table 3.2. * indicates a specimen with unique mutation patterns within the same variant when compared to the other isolates from other specimens. Due to multiple variant infections, some specimens were recorded more than once, therefore the percentages exceed 100%.

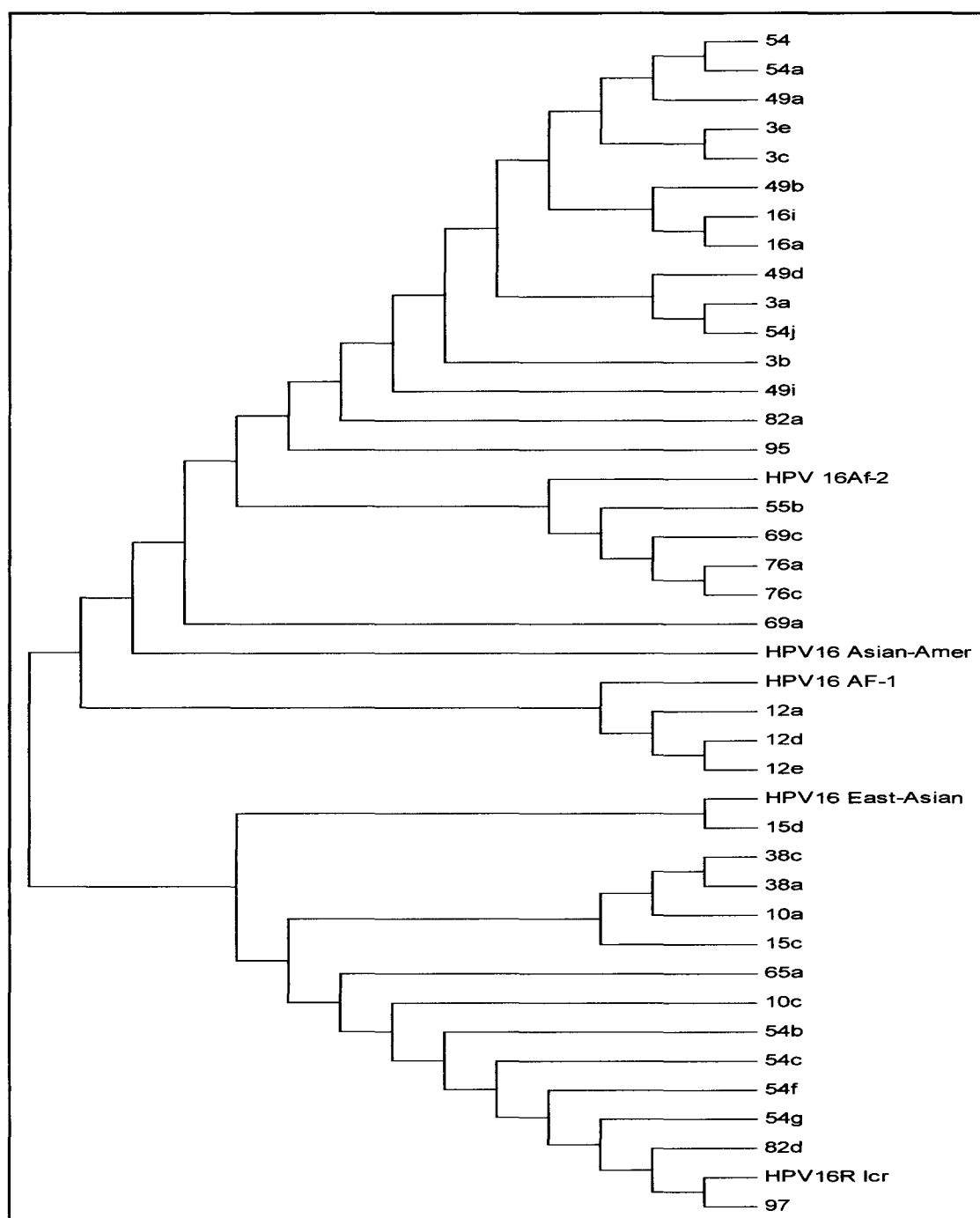


Figure 3.3: Identification of HPV 16 variants by phylogenetic analysis of isolates. The evolutionary history of HPV 16 variants was inferred using the Neighbor-Joining method (Saitou and Nei, 1987), based on the nucleotide sequence of the LCR. The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura *et al.*, 2004). All positions containing gaps and missing data were eliminated from the dataset. Phylogenetic analysis was conducted in MEGA4 (Tamura *et al.*, 2007). HPV 16R lcr, HPV 16 Af-1 and Af-2, HPV 16 East-Asian and HPV 16 Asian-Amer represent GenBank sequences of LCR from the E, Af-1 and Af-2, A and AA variants, respectively. The GenBank accession numbers for the variants were K02718, AF472508, AF472509, AF534061 and AF402678, respectively. The sequences analysed are indicated by specimen number followed by a letter between a-j.

3.4 Discussion

The identification and analysis of HPV 16 variants found in specimens from HIV positive women showed the presence of both E and non-E variants. All the non-E variants were identified; these were the A, AA and Af variants. Most of the observed mutations have been previously reported, with the exception of five nucleotide changes that have not been reported. These were observed in isolates which belong to the Af-1 and Af-2 variants.

Co-infection with multiple HPV 16 variants has been reported in previous variant studies (Stewart *et al.*, 1996) and this was also noted in our study. This suggested sexual mixing behaviour between different populations. Migration of host populations explains the distribution of variants across countries (Stewart *et al.*, 1996; Pande *et al.*, 2008). The presence of both E and non-E variants in our study is evidence of this. The majority of variants in our study were the Af variants (67%, 10/15 specimens), with the Af-2 type accounting for 90% (9/10) of Af variants. In comparison, the E variants were reported in only 47% (7/15) of the specimens. The observed prevalence of variants in this study conforms to the belief that specific variants often cluster in their geographic location and dominate by continent and lineage (Yamada *et al.*, 1997; Bernard *et al.*, 2006). Migration of host, founder effects and co-evolution of HPV may have contributed an important role to effect this distribution (Stewart *et al.*, 1996; Yamada *et al.*, 1997).

Although both E and non-E variants could be identified based on previously reported mutations in the isolates, some uncommon mutations within specific isolates were observed. These often occurred as unique mutations within single specimen, some of which have not been reported previously. An example of this was specimen HH049, an Af-2 variant, which exhibited a unique transversion mutation A7153T in all the four sequenced isolates (Table 3.2). Some uncommon, possibly new mutations within the LCR were observed. These mutations were A7230G, C7550A, T7865C, C7874A, and G59A. The occurrence of these previously unreported mutations may suggest that the point mutations are very frequent within LCR region. As a result, the sequence variations arising from these mutations increase genomic diversity of HPV 16.

Some variants may have high oncogenic potential. The detection of mutations within and upstream of transcription factor binding sites could be responsible for increasing the oncogenic potential of the variants (Pande *et al.*, 2008). A particular example in this study was the commonly observed mutation G7518A (73%, 11/15) in the YY1 binding site, which increases the p97 promoter activity. This increases transcription of the oncoproteins E6 and E7 and potentially increases transformation of cells. This would modify the biological properties of the virus, potentially resulting in cervical disease in individuals harbouring these viruses (Kämmer *et al.*, 2000; Tornesello *et al.*, 2004; Tanzi *et al.*, 2009). This mutation occurred among the Af-2 variants and was absent in the E variants except in one isolate (HH015), which had the most unique mutations.

There are several YY1 binding sites that are affected by the mutations observed in these variants. Those mutations occurring within the YY1 binding site are found in the majority of cervical cancer patients throughout the world. Mutations in this site are absent from non-cancer cases therefore they probably play a major role in carcinogenesis (Alencar *et al.*, 2007; Pande *et al.*, 2008). Other mutations that occur at lower frequency in populations include those in the Oct1 binding site. Nucleotide changes in this region result in the down-regulation of HPV expression (Pande *et al.*, 2008). Although these mutations have been shown to increase disease severity in different ways, the host's genetic component and possibly also HLA haplotypes influence development of invasive cancer.

A mutation in the GRE region (G7385C) was one of the common sites of variation within the LCR as it was detected in 60% (9/15) of the specimens, all of which are Af-2 variants. The occurrence of GRE/PRE within the LCR has a significant impact on the ability of HPV to transform cells and result in malignancy. The discovery of these regulatory elements showed the critical role played by the LCR in determining the outcome of an HPV infection (Chan *et al.*, 1989). Burk and colleagues (2003) observed that mutations in the GRE region result in HPV 16 variants responding differently to hormonal changes and impacting differently on oncogenicity of variants. The mechanisms by which this occurs still remain to be elucidated (Burk *et al.*, 2003). A mutation in the AP1 binding site (A7633C) within the Af-2 variants was also observed and is known to down-regulate HPV expression (Pande *et al.*, 2008). Most of the mutations whose effects are well characterised and known to increase risk of cervical

disease/cancer were observed among the Af variants. This may suggest that the claims that Af variants are associated with disease may be true.

Due to lack of clinical data, it was not possible to further analyze the particular variants in relation to disease severity among the HIV infected women as the disease status on specific individuals with HPV 16 was not available. However, it was clear that HPV 16 was one of the most common types among women with abnormal Pap smears (29%) (Moodley *et al*, 2009). Although there are claims that Af variants are associated with increased risk of developing cervical neoplasia/cancer, Tu and colleagues (2006) reported that the Af variants were not associated with increased risk of neoplasia. Their argument was based on results from their study on HPV 16 variants in South African women whose HIV status was unknown (Tu *et al.*, 2006). Therefore, such an investigation would prove helpful in understanding the risk associated with HPV variants affecting South African women infected with HIV.

In this study, the variants were determined from polymorphism of the LCR since it is considered to be the best region for variant identification (Stewart *et al.*, 1996). The high level of sequence variation reported in this study proves high genomic diversity among HPV 16 genome and it is possible that this pattern may be shared by other genotypes that may be less prevalent in populations. Similar sequence variations observed in HPV 39 sequences may suggest that more HPV types have similar variation patterns.

Chapter 4

An HPV-specific metagenomic analysis on one HIV infected specimen

4	INTRODUCTION.....	65
4.1	MATERIALS AND METHODS.....	66
4.1.1	<i>Rolling-circle amplification of DNA for pyrosequencing</i>	<i>66</i>
4.1.2	<i>HPV genotyping by pyrosequencing</i>	<i>68</i>
4.1.3	<i>Restriction enzyme analysis and cloning of HPV genome to detect recombination.....</i>	<i>70</i>
4.1.4	<i>Sequencing and analysis</i>	<i>71</i>
4.2	RESULTS.....	71
4.2.1	<i>Rolling circle amplification of DNA.....</i>	<i>71</i>
4.2.2	<i>HPV detection by pyrosequencing</i>	<i>73</i>
4.2.3	<i>Detection of recombination in HPV.....</i>	<i>78</i>
4.3	DISCUSSION.....	84

4 Introduction

In Chapter 2, HIV infected women were found to display a wide spectrum of HPV types. The observed diversity of HPV types may suggest that there could be more mucosal HPV types that could still be identified if a different approach is used. When considering the high number of multiple infections observed, it is possible to conclude that the detection method used in Chapter 2 may not include as many mucosal HPV types as there exist in HIV infected women, due to HIV positive women being known to have higher rates of HPV infection compared to their HIV negative counterparts (Banura *et al.*, 2008). Therefore, more HPV types are likely to infect these women than those incorporated in and detected by the Roche Linear Array HPV genotyping test.

In addition to the PCR-based detection methods, other methods that have been utilised in HPV genotyping include pyrosequencing. HPV genotyping of clinical specimens through pyrosequencing of PCR products with multiple HPV-specific and general primers has been successful (Gharizadeh *et al.*, 2001; Gharizadeh *et al.*, 2005). In this chapter, the use of pyrosequencing technique on genomic DNA to detect HPV types was explored. Pyrosequencing was performed following rolling-circle DNA amplification (RCA) of circular HPV DNA. The multiply-primed sequence non-specific RCA procedure was used together with pyrosequencing method to detect HPV types in one of the clinical specimens.

Rolling-circle replication is a replication mechanism used by certain circular DNA genomes such as plasmids and viruses (Rector *et al.*, 2004). During rolling-circle replication (RCR) the DNA is unwound and DNA polymerase synthesises a sequence complementary to the template DNA strand. This process can be mimicked *in vitro*, for a technique known as RCA, through the use of random hexamers and bacteriophage Φ 29 DNA polymerases (Rector *et al.*, 2004). RCA has been applied in the identification of novel types or strains of viruses with circular DNA genomes. The discovery of novel viruses used to be done through Southern blot hybridisation, which allowed identification based on homology between related sequences. This means that there had to be prior knowledge of the sequence of at least one virus type. However, this only allowed determination of part of the genome (Rector *et al.*, 2004). The complete genome was then

elucidated by an inverse PCR method. Alternatively, amplified genomic DNA was digested with restriction enzymes and cloned before sequencing (Rector *et al.*, 2004). When the RCA method was discovered, genomic DNA could be analysed for new divergent viruses in a sequence non-specific manner by directly sequencing RCA products. The ability of RCA reactions to maintain sequence specificity of DNA during amplification led to their application in genotyping and in mutation detection studies (Demidov, 2005). Therefore, RCA was used in our study to amplify HPV genomes prior to sequencing to detect the HPV genotypes in one specimen.

In this study, RCA coupled with pyrosequencing was used for genotyping HPV in an individual co-infected with HIV. The results were compared to those of the Roche Linear Array HPV Genotyping test for the same specimen, HH015, with 12 HPV types. This specimen was chosen as it had the highest multiple HPV infection (as detected by the genotyping test) compared to the others. DNA from that specimen only was pyrosequenced due to time and budget limitations. The specimen was investigated for other existing but undetected, new and recombinant HPV types. Recombination was investigated by cloning and sequencing of complete HPV genome. The persistence of multiple HPVs can promote recombination between the different HPV types. South Africa has a high prevalence rate of HIV and with increased HPV infections in these women, investigation of the possibility of recombination events in HPV genomes among South African women was necessary.

4.1 Materials and Methods

4.1.1 Rolling-circle amplification of DNA for pyrosequencing

HPV circular DNA from one sample infected with 12 HPV types and co-infected with HIV (HH015) was amplified by RCA using the Illustra™ TempliPhi 100 Kit (Amersham Biosciences, GE Healthcare, UK). The reaction was performed according to manufacturer's instructions. One microlitre of genomic DNA was mixed with 5µl of sample buffer in a 0.2ml tube. The sample buffer contained exonuclease-protected random hexamers. The mixture was heated at 95°C for 3 minutes to denature the ds DNA and cooled on ice. A premix of the TempliPhi enzyme mix (0.2µl) and TempliPhi reaction buffer (5µl) was prepared in a separate tube. The enzyme mix contained the Φ29 DNA polymerase, extra dNTPs and extra hexamers in 50% glycerol. The TempliPhi

reaction buffer contained dNTPs and salts for enzyme activity. Five microlitres of the premix were added to the cooled sample and mixed by vortexing. The reaction was incubated at 30°C for 18 hours on an AB9700 machine (Applied Biosystems). Amplification was initiated by the hybridisation of exonuclease-protected random hexamer primers to the circular DNA. The procedure continued in a chain reaction where the kinetics of multiple hybridisation, primer-extension and strand displacement occurred as illustrated in Figure 4.1. Due to the strong displacement capability of bacteriophage Φ 29 DNA polymerase, strand displacement occurred when the enzyme reached a downstream extended primer (Rector *et al.*, 2005; Demidov, 2005). This allowed for amplification of the nucleotide sequences by several magnitudes, yielding several complementary tandem repeats of an original DNA circle. The result was a distinct set of concatemers of ds HPV DNA. After the reaction, the mixture was incubated at 65°C for 10 minutes to inactivate bacteriophage Φ 29 DNA polymerase. Two microlitres of sample were electrophoresed on 0.8% agarose gel as outlined in Appendix Aiii. The rest of sample was stored at -20°C until further use in subsequent reactions.

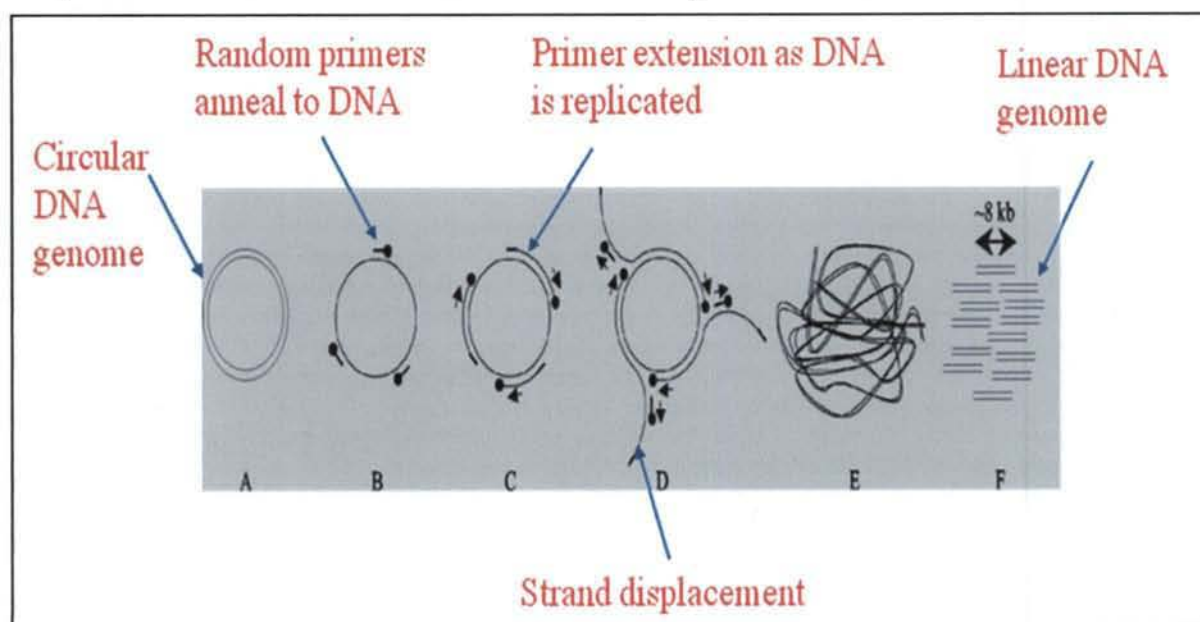


Figure 4.1 A-F: Schematic diagram of rolling-circle amplification in circular genomes. Random hexamers annealed to DNA template at multiple sites (B). The enzyme extended each of the primers until it reaches another downstream extended primer (C). Bacteriophage Φ 29 DNA polymerase allowed strand displacement to occur (D). The displaced single strand acts as an available template and more hexamers annealed to it for further extension (E). Continued extension and strand displacement resulted in amplification to give double-stranded DNA (E). Restriction endonuclease digestion using enzymes with single recognition site within the circular DNA genomes yields linear ds HPV DNA (F) (modified from Rector *et al.*, 2004).

4.1.2 HPV genotyping by pyrosequencing

This technique was used to detect HPV genotypes in one specimen, HH015. The genomic DNA was randomly fragmented by nebulisation, to create a library of between 300-800bp fragments and short adaptor molecules were attached to the DNA library (Figure 4.2).

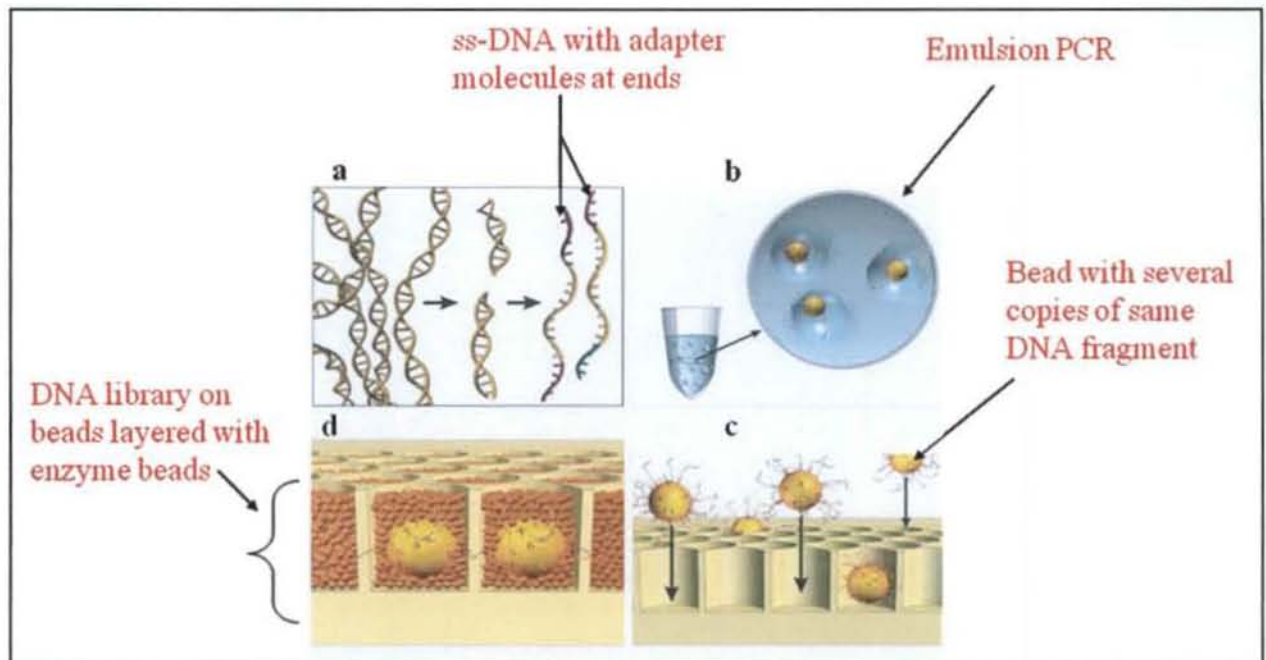


Figure 4.2: Sample preparation for pyrosequencing on the 454 Genome Sequencer (GS) FLX™ system (Roche Diagnostics). DNA was randomly fragmented to create a library of between 300-800bp fragments. Adaptor molecules were attached to the single-stranded fragments generated by fragmentation of genomic DNA and bound to beads (a). The beads were captured in droplets of a PCR-reaction-mixture-in-oil emulsion before clonal amplification (b). Each library bead was loaded into a single well (c) and DNA library beads were layered with enzyme beads (d) before being loaded on a 454 PicoTiterPlate device (Modified from Margulies *et al.*, 2005).

The adaptor molecules were used in subsequent purification, amplification and sequencing reactions (Droege and Hill, 2008). The DNA fragments were captured onto beads under conditions promoting only a single-stranded DNA molecule per bead, and clonally amplified. This generated millions of unique DNA templates per bead as shown in Figure 4.2c (Margulies *et al.*, 2005; Droege and Hill, 2008). The generated DNA library beads were added to the DNA Bead Incubation Mix containing exonuclease deficient (exo^-) Klenow DNA polymerase I (DNA polymerase I). Each bead was

deposited by centrifugation into a 44 μ m picotitre well on a picotiter device. The library beads were layered with Enzyme Beads (with sulfurylase and luciferase enzymes) onto a picotiter device. The device was mounted onto the GS FLX™ instrument where sequencing was performed in a number of cycles (Droege and Hill, 2008; Margulies *et al.*, 2005). During DNA sequencing, a single base was incorporated by DNA polymerase I to extend the bound DNA template (Figure 4.3). The synthesis of each base resulted in the release of an inorganic pyrophosphate (PPi) during each nucleotide flow cycle. The PPi was converted to ATP by sulfurylase enzyme in a reaction that generated enough energy for oxidation of luciferin by luciferase (Figure 4.3). The luciferin produced light that was quantitated by a charge-coupled device (CCD) camera to determine the number of bases incorporated (Bentley, 2006; Ronaghi, 2001). The camera simultaneously captured images from different wells that contained template-carrying beads where a nucleotide had been incorporated in each cycle of nucleotide flow. Each cycle of nucleotide flow was followed by a wash with apyrase enzyme to remove all unincorporated nucleotides before the next flow started. This generated several base reads which were analysed on the *de novo* assembler (Margulies *et al.*, 2005).

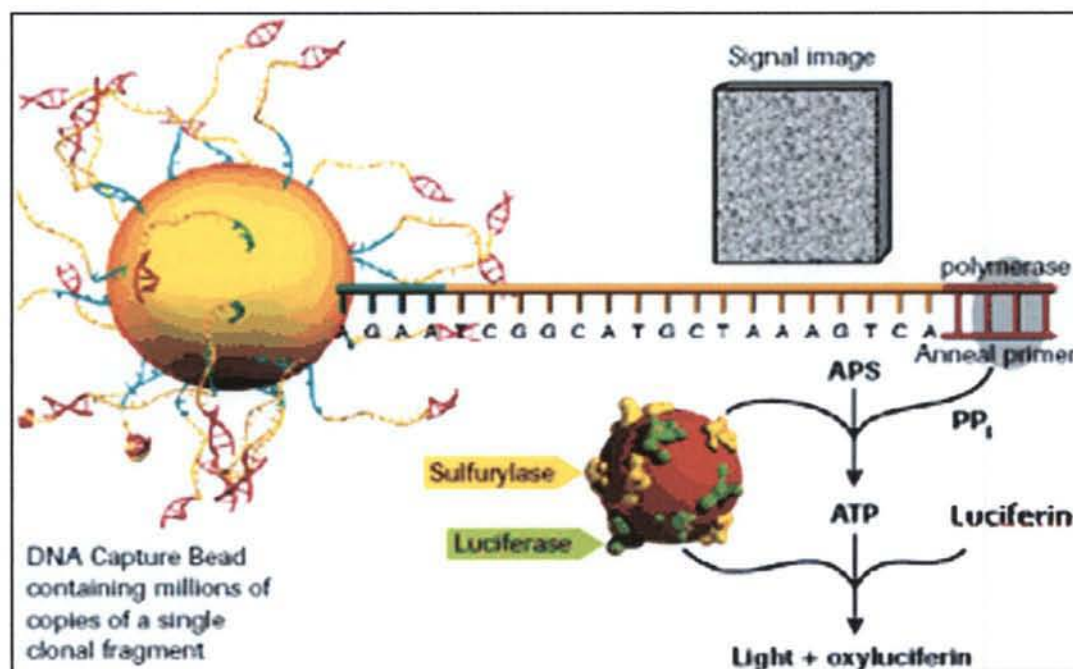


Figure 4.3: Pyrosequencing using the 454 sequencing technology of sequencing-by-synthesis. The four nucleotide bases, ACGT, were cycled several times through the sequencer and each base was added one at a time to the template, in yellow. The adapter molecules are shown in green. A chemiluminescent signal was generated during the cycle and processed to determine base sequence (Taken from Droege and Hill, 2008).

Pyrosequencing was performed at Inqaba Biotech, Pretoria. The sequence results were assembled using the *de novo* sequence assembler, NewblerAssembler at Inqaba Biotech. The assembler showed overlapping regions of the same HPV types by alignment of generated reads. This allowed for the creation of overlaps between reads and subsequently larger sequences from overlapping short sequences. Assembled sequences were searched against the HPV database using the Basic Local Alignment Search Tool (BLASTN) from the National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov). The sequences that did not result in known HPV sequences by BLASTN were also searched for potential HPV ORFs, using BLASTX.

4.1.3 Restriction enzyme analysis and cloning of HPV genome to detect recombination

The restriction enzyme analysis was performed in order to facilitate cloning and subsequent sequencing of full length HPV genome and detect recombination events. Circular DNA from HH015 amplified by rolling-circle mechanism was analysed with restriction enzymes *EcoRI*, *BglI* and *BamHI* to identify the restriction patterns for each enzyme. The enzymes were selected based theoretical analysis using DNAMAN software (version 5.2.9) to identify enzymes with one restriction site within full length genomes of HPV types detected in sample HH015. The RCA product was analysed with the mentioned restriction enzymes as outlined in Appendix Aiv.

The amplified DNA was digested with *EcoRI* as outlined in Appendix Aiv. The digestion mix was ligated into a dephosphorylated *EcoRI* digested pUC19 vector (Appendix Avi). The ligation products were transformed into *E. coli* DH5 α cells as in Appendix Avii. The transformed cells were plated onto Luria-Bertani agar (LA) plates with ampicillin for selection of clones. Single colonies of *E. coli* from LA plates were inoculated into Luria-Bertani broths (LB) to replicate cloned DNA. The LA plates and LB were prepared as in Appendix Ai and all growth of *E. coli* was at 37°C for 18 hours. Bacterial cultures were incubated at the same temperature on a shaker. DNA isolation from bacterial cultures was performed as in Appendix Aii.

4.1.4 Sequencing and analysis

The cloned genome was sequenced using the Sanger chain termination method (Appendix Aviii) at Macrogen, Korea. This was done to identify recombinant regions within the HPV genome. Sequencing was achieved by primer-walking, initially using standard M13 primers shown in Appendix Bii. Subsequent primers were designed from the generated sequence and used for sequencing (Appendix C). This was repeated until the whole genome was sequenced. The sequenced regions were analysed using DNAMAN software version 5.2.9 and searched in the NCBI database using BLAST facility (www.ncbi.nlm.nih.gov/blast). The full genome was analysed for recombination.

4.2 Results

4.2.1 Rolling circle amplification of DNA

Circular DNA amplification by RCA was performed to non-specifically increase circular DNA concentration for subsequent pyrosequencing reactions. DNA was amplified by extension of several hexamer primers, hybridised to ssDNA molecules, by Φ 29 DNA polymerases. This resulted in linear amplification of DNA to give several thousands of copies of the original DNA in solution. The amplification process enriched the DNA sample for the pyrosequencing reaction by increasing the copy number of the different HPV types in HH015 sample. Figure 4.4 shows the agarose gel electrophoresis of RCA product. The RCA product was further analysed by sequencing on the 454 GS FLX™ and the Sanger chain termination method.

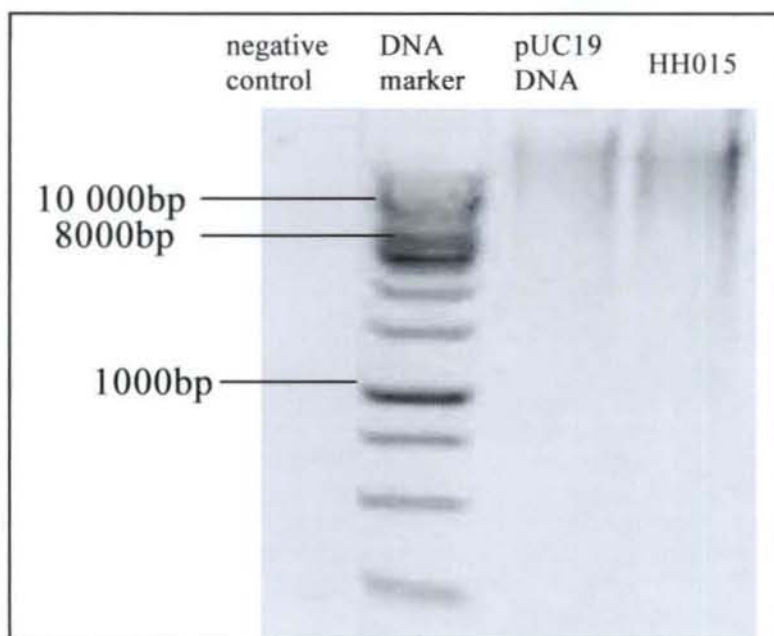


Figure 4.4: Gel electrophoresis of rolling-circle amplification product of DNA from HH015 on 0.8% agarose gel. The plasmid, pUC19, DNA was used as a positive control in the reaction and the negative control was water. Amplified DNA was electrophoresed as a smear, lanes HH015 and pUC19.

Pyrosequencing and the Roche Linear Array HPV genotyping test were used to compare the HPV types identified by the two methods. The Roche Linear Array HPV Genotyping test identified 12 HPV types, shown in Figure 4.5. Pyrosequencing identified seven types; HPV 16, 30, 39, 40, 56, 74 and 86.

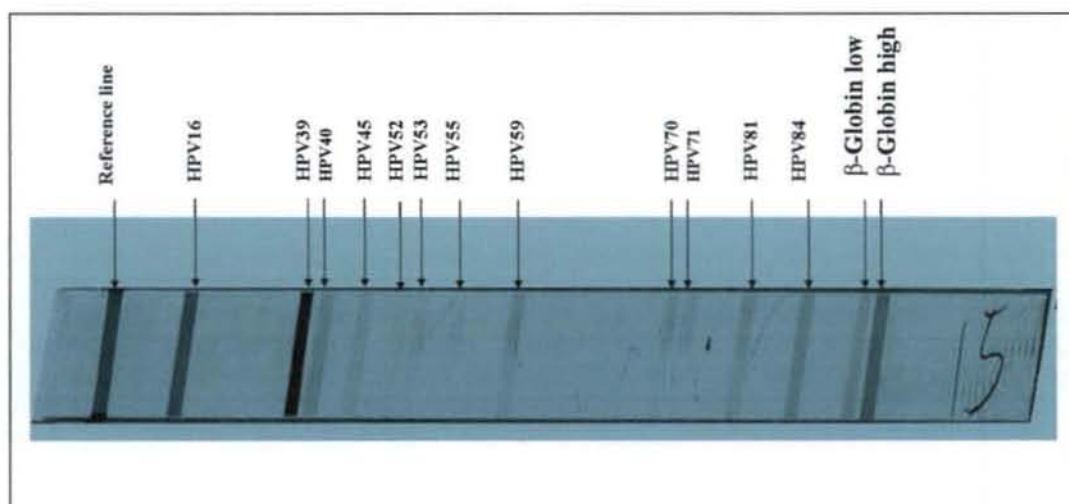


Figure 4.5: The Roche Linear Array HPV Genotyping test on specimen HH015. The HPV types identified in the specimen are shown above. The darker bands indicate higher viral load in comparison with other bands. HPV 39 DNA appeared to be the most abundant in the specimen.

4.2.2 HPV detection by pyrosequencing

Pyrosequencing on DNA from specimen HH015 was performed to detect HPV types in specimen and compare to the types detected by the Roche Linear Array HPV genotyping test. A total of 13128 reads were generated from pyrosequencing. Each read consisted of between 200-300 bases. The reads were assembled into 976 sequences. A summary of the data from the NCBI database search on all the 13128 reads is shown in Figure 4.6. The reads showed homology to DNA from the following organisms; humans, HPV, bacteria, *Pan troglodytes*, *Pongo abelii*, *Mus musculus* and *Macaca mulata* among others. The number of reads for HPV was only 0.19% of the total reads. The majority of the reads had high homology to human DNA. Therefore, most of the reads belonged to human DNA sequences, with 1.21% of total 13 128 reads.

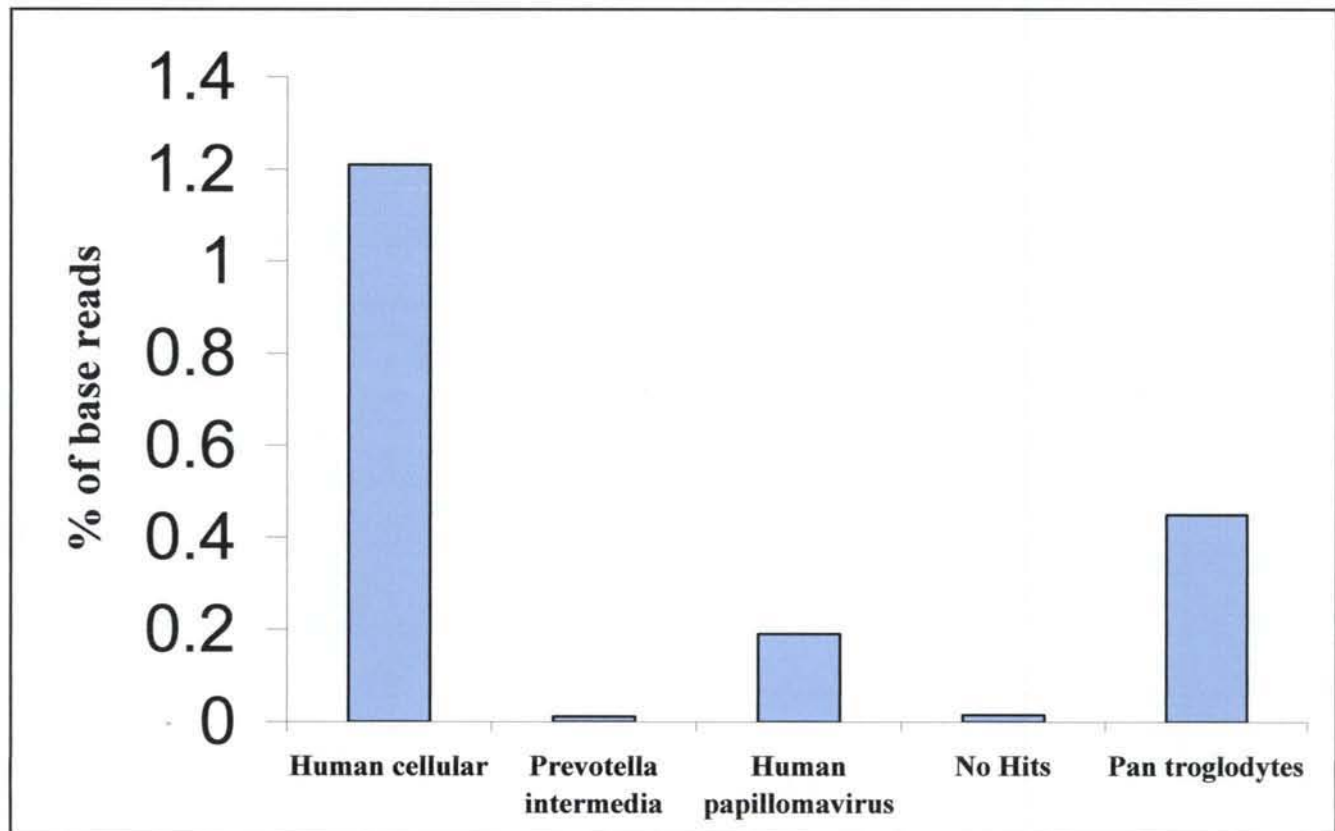


Figure 4.6: Summary of significant BLAST search results of base reads generated by pyrosequencing using the 454 GS FLX™ system, from the NCBI database. The search was conducted by using BLAST on all the reads, n=13128. Most of the reads represented DNA of human origin.

A total of 976 sequences were obtained from the *de novo* assembly of all 13128 reads generated from the pyrosequencing. The sequences were searched in the NCBI database

using BLASTN and 23 (2%) sequences were found to be from known HPV types. The specific HPV type for each of the 24 sequences is listed in Table 4.1. The other organisms which were identified from the database search of the sequences include *Prevotella intermedia*, or a relative: this is a bacterium infecting the mouth and is responsible for periodontitis (Finegold, 1995), and this was identified from one sequence of length 3087 bases. 85% of this sequence (query coverage) had 85% homology to the bacterial sequence. This significant homology result suggests the presence of this organism in the total DNA extracted from HH015. *Bacteroides fragilis* nucleotide sequence was also identified, showing 88% query coverage and 72% homology. Three other sequences (2 x 241 and 133 bases) showed high homology (100%) to a Human endogenous retrovirus K (HERV K), a normal part of the human genome in most individuals and does not cause disease (Boller *et al.*, 1993).

Some of the sequences had high homology (97%) across the entire 241 base sequence (100% query coverage) to the pan troglodytes DNA when searched in the NCBI database. These were considered to be sequence similarities between human DNA and *Pan troglodytes* DNA since humans and chimpanzees share common ancestry. Most of the other organisms were identified from sequence lengths of 100-400bp. The sequences were short and their coverage in the database varied from 8-88% from *Mus musculus* and *Macaca mulata*. These were considered to be related to human DNA but not present in sample. Several other sequences (48/976, 5%) had no significant homology to any organisms in the database. The size of these sequences ranged from 98-1356 bases, with a median length of 241 bases. No new HPV types were detected and obviously recombinant HPV sequences were absent. All the sequences that did not yield significant homology (5%, 48/976) from the nucleotide database were searched against the protein database in NCBI for potential HPV ORFs. Among these sequences, no HPV proteins were detected from the database search.

The seven HPV types identified were 16, 30, 39, 40, 56, 74 and 86. While some HPV types were identified from one sequence, other HPV types identified were represented by more than one sequence, as shown by the sequenced regions in Figure 4.7. The coverage of the sequences was between 98-100% (Table 4.1) except for two sequences, one for HPV 16 and HPV 74, which had 88% and 71% coverage, respectively. This was probably a result of the programme used to assemble sequences. The similarity/homology

of all nucleotide sequences to regions of their respective HPVs in the database ranged from 97-100% (Table 4.1).

The number of full-length genome sequences (~ 7.9kb) of HPV from this sequencing was limited to two. These were HPV 39 and HPV 40 (Figure 4.7d and e). The other HPV genomes were fragmented, with long regions missing, for example HPV 30 (Figure 4.7f).

Table 4.1: Summary of HPV sequence identity from BLASTN**(www.ncbi.nlm.nih.gov/blast)**

Length (bp)	HPV Identity	GenBank accession no.	Query cover (%)	% identity	Region in genome
208	HPV 30	X74474	99	99	7388-7183
192	HPV 30	X74474	100	98	487-677
137	HPV 30	X74474	100	100	5839-5703
334	HPV 16	K02718	100	100	315-648
145	HPV 16	K02718	100	97	245-389
950	HPV 16	K02718	88	99	7904-7172
2229	HPV 16	K02718	99	99	644-2872
2744	HPV 16	K02718	100	99	2860-5606
1505	HPV 16	K02718	100	99	5618-7124
241	HPV 56	X74483	100	98	5840-5600
184	HPV 56	X74483	100	100	4358-4541
221	HPV 56	EF177181	99	99	1523-1303
184	HPV 56	X74483	99	99	1522-1302
228	HPV 56	EF177179	100	100	7268-7496
	HPV 56	X74483	92	100	7280-7491
231	HPV 74	AF436130	99	97	699-473
	HPV 74	U40822	99	96	3649-3420
218	HPV 74	U40822 (cnds)	100	99	1354-1136
	HPV 74	AF436130	100	95	6285-6067
672	HPV 74	AF436130	99	98	5597-6267
	HPV 74	U40822	99	94	666-1336
122	HPV 74	AF436130	100	99	2800-2921
419	HPV 74	AF436130	98	98	6624-7037
	HPV 74	U40822	98	93	1692-2103
241	HPV 74	AF436130	71	97	322-493
241	HPV 86	AF349909	100	96	7983-7777, 32-1
7831	HPV 39	M62849	100	99	7833-2673, 2672-1
1896	HPV 40	X74478	100	99	2738-839
6058	HPV 40	X74478	100	99	7909-2764, 914-1

HPV sequences as identified from the NCBI database through BLAST. HPV 39 was the only HPV type with a continuous whole genome sequence. The two HPV 40 sequences give an almost full genome sequence, with 73 bases missing to complete the genome. HPV 16 had six fragmented sequences, all of which show that they were from different variants of HPV 16 as indicated by the GenBank accession numbers. HPV 56 had five sequences, all from different regions of this type. HPV 74 had six different sequences of varying lengths, which belong to either the HPV 74 subtype AE10 (AF436130) or the HPV 74 coding domain sequence (U40822). Only one 241 base sequence resulted in HPV cand86 (cand refers to candidate HPV type, which was cloned and characterised from PCR products (de Villiers *et al.*, 2004). This will be referred to as HPV 86). Query cover defines the length (%) of the sequence being searched in database against the resulting sequence from the NCBI database. For example, 99% query cover for 208 bases sequence indicates that 99% of 208 bases were aligned to the sequence found in database, in this case, HPV 30.

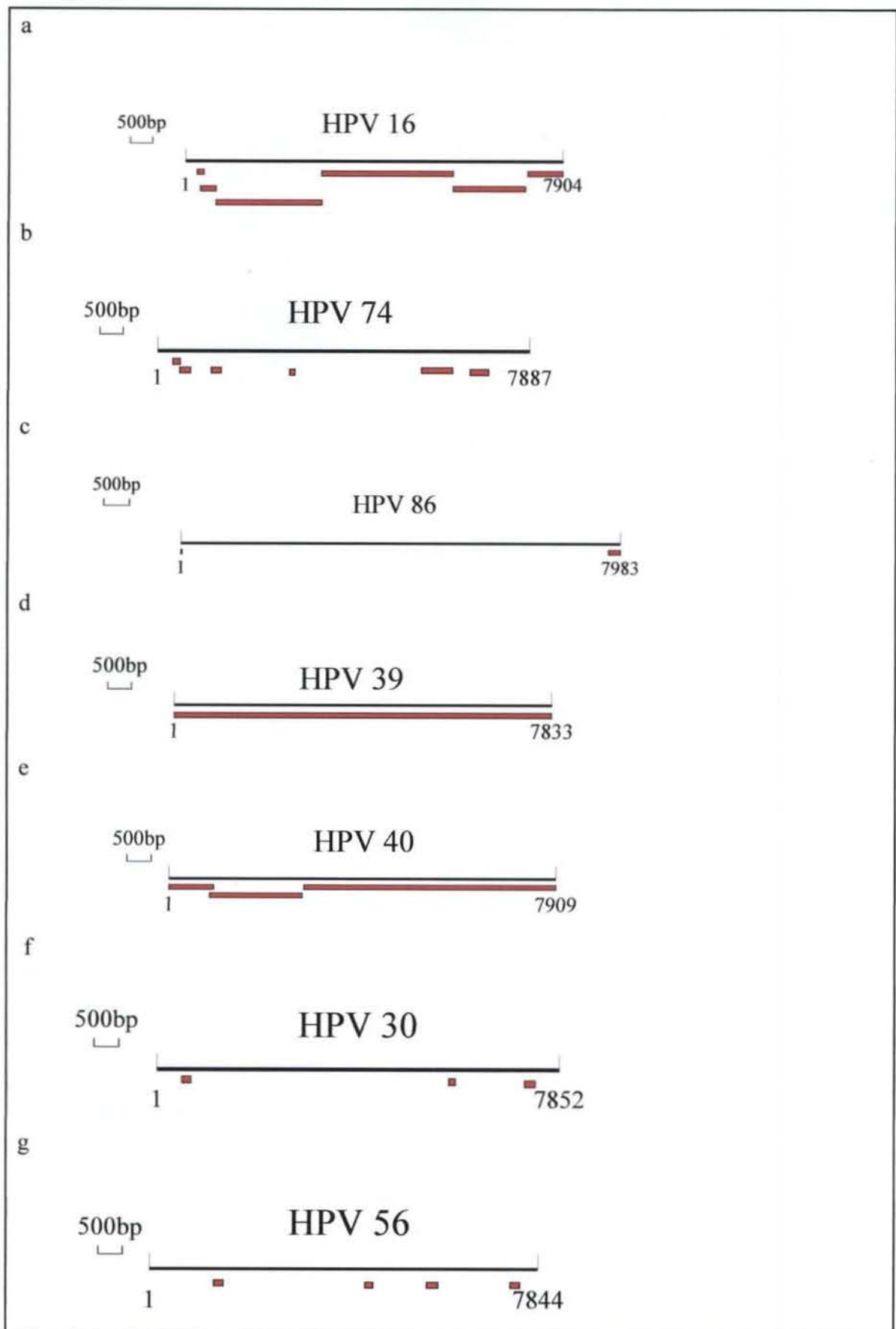


Figure 4.7: A schematic representation of the sequenced regions of the HPV types identified by pyrosequencing.

Some of the HPV types detected by pyrosequencing were not detected by the Roche Linear Array HPV Genotyping test. The HPV types detected by each of these methods are listed in Table 4.2 below, highlighting the differences and similarities in the types identified by each method.

Table 4.2: Detection methods used for HPV types in specimen HH015 and the types identified in each method.

Detection method	HPV types detected
Roche Linear Array HPV genotyping test	16, 39, 40, 45, 52, 53, 55, 59, 70, 71, 81, 84
Pyrosequencing with the 454 Genome Sequencing FLX™ system	16, 30, 39, 40, 56, 74, 86

The types common in both methods are highlighted in red.

4.2.3 Detection of recombination in HPV

Restriction enzyme analysis was performed on the RCA product of DNA from specimen HH015 to facilitate cloning of complete linearised HPV genomes. This specimen had 12 different HPV types (Figure 4.5). The restriction enzymes that digested the HPV types present in HH015 were *BglI*, *BamHI* and *EcoRI*. The types digested by each of these enzymes are indicated in Table 4.3. The restriction pattern for each enzyme is illustrated in Figure 4.8.

Table 4.3: Restriction enzymes that digested HPV types present in specimen HH015

HPV type	Restriction enzyme		
	No. of enzyme sites and enzyme digestion sites		
	<i>EcoRI</i>	<i>BamHI</i>	<i>BglI</i>
16	(2) 6818, 7453	(1) 6150	
30^a	(1) 1483	(1) 6511	
39	(1) 6824	(2) 1473, 4738	(1) 6572
40		(3) 6661, 6957, 7089	(1) 4409
45		(2) 920, 6933	
52	(2) 111, 5093	(1) 7559	
53	(1) 6830	(1) 6098	
55	(2) 37, 2170	(2) 574, 5032	
56^a	(3) 2360, 7449, 7491	(1) 5521	
59		(1) 69	
70	(1) 3849	(2) 5847, 6900	
74^a		(1) 4920	(1) 6511
81		(1) 4758	(4) 1649, 2086, 2948, 6796
84	(1) 7888		(1) 4465
86^a		(3) 250, 6519, 6983	(1) 4443

^a – HPV types identified by sequencing on 454 GS FLX™ system but not the Roche Linear Array HPV Genotyping test. HPV 71 was not digested by the selected enzymes.

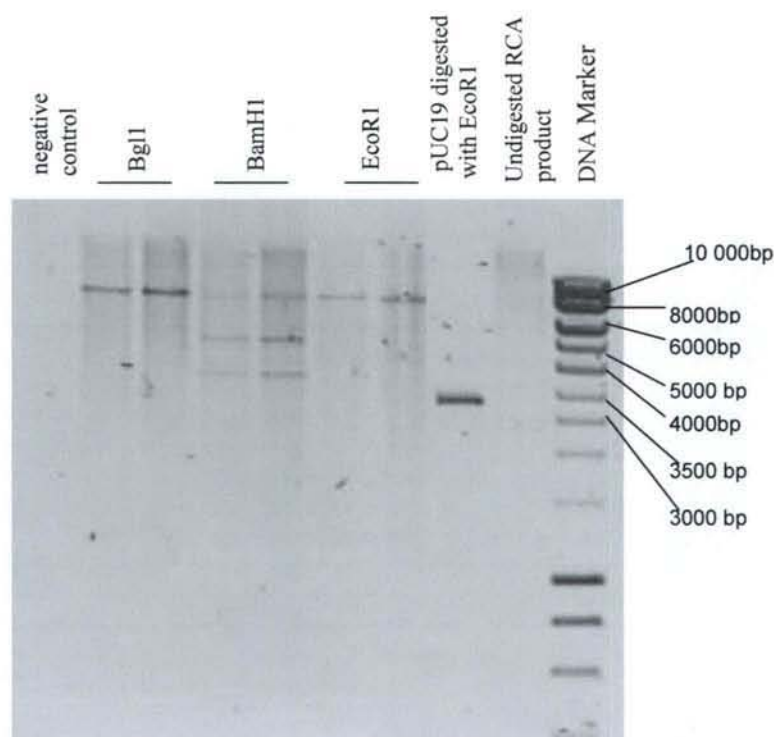


Figure 4.8: Gel electrophoresis of RCA product of DNA from specimen HH015. The RCA product migrated through the agarose gel as a smear (undigested RCA product). The DNA was digested with *BglI*, *BamHI* and *EcoRI*. The negative control was a no DNA sample. The positive control, pUC19, was digested with *EcoRI*. Digestion of RCA product with *BglI* and *EcoRI* gave one fragment-8000bp, while restriction analysis with *BamHI* gave three fragments: 8000bp, 4200bp and 3800bp.

The enzyme *EcoRI* was selected for the cloning of complete HPV genomes for the detection of recombination based on restriction pattern observed when DNA was digested (Figure 4.8). *EcoRI* generated one 8000bp size fragment and digested more types compared to *BglI*. Digestion of DNA with *BamHI* resulted in three fragments (Figure 4.8). According to the digestion pattern in Table 4.3, *EcoRI* digested less HPV types in specimen HH015 than *BamHI* (Table 4.3). The digestion pattern for *EcoRI* shown in Figure 4.8 showed that HPV 16, 52, 55 and 56 did not show any bands as expected since they have more than one *EcoRI* site (Table 4.3). HPV 81 DNA digested with *BglI* did not show any bands in the agarose gel (Figure 4.8). The *EcoRI* digested RCA product was cloned into pUC19 vector and two clones (15.2 and 15.3) were selected for sequencing to identify regions of recombination in HPV genome.

The *EcoRI* linearised and cloned HPV genome was initially sequenced with M13 forward and reverse primers. Analyses of the sequences from both clones (15.2 and 15.3) identified the cloned genome as HPV 39 (GenBank accession number M62849), as shown in Figure 4.9 and 4.10. After the complete genome was sequenced (Figure 4.11), the short sequences were aligned to give one complete HPV 39 sequence. All the

sequences had between 98-100% homology to HPV 39. No recombination regions were detected although several point mutations were present within the sequence (Figure 4.12). The analysis of the same DNA sample using a different restriction enzyme could not be completed within the time frame of this project, hence no further clones could be sequenced.

```
>gb|M62849.1|PPHT39 Human papillomavirus ORFs
Length=7833

Score = 1153 bits (1278), Expect = 0.0
Identities = 641/642 (99%), Gaps = 0/642 (0%)
Strand=Plus/Plus

Query 1      ATGTCAAAGGATGCTCCAGCACCTGAAAAGAAAGATCCATATGACGGTCTAAAGTTTGG 60
            |||
Sbjct 6923    ATGTCAAAGGATGCTCCAGCACCTGAAAAGAAAGATCCATATGACGGTCTAAAGTTTGG 6982

Query 61      GAATGTTGACTTAAGGGAAAAGTTTAGTTTGGAACTTGATCAATTCCTTTGGGACGTAA 120
            |||
Sbjct 6983    GAATGTTGACTTAAGGGAAAAGTTTAGTTTGGAACTTGATCAATTCCTTTGGGACGTAA 7042

Query 121     ATTTTGTTGCAGGCCAGGGTCCGCGAGCGGCCCTACTATAGGTCCCCGAAAGCGGCGTGC 180
            |||
Sbjct 7043    ATTTTGTTGCAGGCCAGGGTCCGCGAGCGGCCCTACTATAGGTCCCCGAAAGCGGCGTGC 7102

Query 181     TGCATCCACTTCCTCGTCTCAGCTACTAACAACAAACGTAAACGTGTGTCTAAATAATG 240
            |||
Sbjct 7103    TGCATCCACTTCCTCGTCTCAGCTACTAACAACAAACGTAAACGTGTGTCTAAATAATG 7162

Query 241     catgtgtatgcttctgttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
            |||
Sbjct 7163    CATGTGTATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7222

Query 301     tgtgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
            |||
Sbjct 7223    TGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7282

Query 361     CATGTGTGATTGCACACCTGTGACTAACAGTGTATTGTGTGTGTGTGTGTGTGTGTGTGT 420
            |||
Sbjct 7283    CATGTGTGATTGCACACCTGTGACTAACAGTGTATTGTGTGTGTGTGTGTGTGTGTGTGT 7342

Query 421     AACATTTTATACATAATCTATATGCCCTACCTAAGGTGTGTGTGTGTGTGTGTGTGTGTGT 480
            |||
Sbjct 7343    AACATTTTATACATAATCTATATGCCCTACCTAAGGTGTGTGTGTGTGTGTGTGTGTGTGT 7402

Query 481     TTTTACATTGTTGTATGCGTTTCTACATTTTATACTTCGCCATTTTGTGGCGACCGAAG 540
            |||
Sbjct 7403    TTTTACATTGTTGTATGCGTTTCTACATTTTATACTTCGCCATTTTGTGGCGACCGAAG 7462

Query 541     TCGGTCGTGGTTGAGCAATTTTAACTAGTGGAAACACCTTTCTCAGCAAAAACA 600
            |||
Sbjct 7463    TCGGTCGTGGTTGAGCAATTTTAACTAGTGGAAACACCTTTCTCAGCAAAAACA 7522

Query 601     TGTCTTTACCTTAGGTTTCACTGTCATAGTTGGCACTGGTAA 642
            |||
Sbjct 7523    TGTCTTTACCTTAGGTTTCACTGTCATAGTTGGCACTGGTAA 7564
```

Figure 4.9: BLASTN search result from a sequence generated by M13 forward primer, showing 100% homology to HPV 39. Subject represents Human papillomavirus ORFs sequence, also known as HPV39 and the query sequence is shown in the top line.

```

>|gb|M62849.1|HPV39 Human papillomavirus ORFs
Length=7833

Score = 1561 bits (1730), Expect = 0.0
Identities = 865/865 (100%), Gaps = 0/865 (0%)
Strand=Plus/Minus

Query 54 GAATTCATAGTGTGAATATAAGACATAACATCAGTTGTTAATGTGACAGTACACAGTTGA 113
|
Sbjct 6829 GAATTCATAGTGTGAATATAAGACATAACATCAGTTGTTAATGTGACAGTACACAGTTGA 6770

Query 114 AATATAAATTGTAAATCATACTCCTCCACGTGCTGGTATATTCCTTAACTTAGAAGGA 173
|
Sbjct 6769 AATATAAATTGTAAATCATACTCCTCCACGTGCTGGTATATTCCTTAACTTAGAAGGA 6710

Query 174 TCATATGTAGAGGTATGGAGACTCTATAGAGGTAGATAATGTAAAGTTGGTACTACGG 233
|
Sbjct 6709 TCATATGTAGAGGTATGGAGACTCTATAGAGGTAGATAATGTAAAGTTGGTACTACGG 6650

Query 234 GTAGTGCCCAACAGTAAGAAATAATTGATTATGCCAACATATACCAITGTTGTGGCCC 293
|
Sbjct 6649 GTAGTGCCCAACAGTAAGAAATAATTGATTATGCCAACATATACCAITGTTGTGGCCC 6590

Query 294 TGGGCCCTTATGTAGCCAAATAGGCTTATTAATAACTGGGAATCAGAGGTTACCATGGA 353
|
Sbjct 6589 TGGGCCCTTATGTAGCCAAATAGGCTTATTAATAACTGGGAATCAGAGGTTACCATGGA 6530

Query 354 CCGCTGGGAGAGGGGAGTACAGAACTACCGGGGTTTGACGTATATCTGTGCCCTTA 413
|
Sbjct 6529 CCGCTGGGAGAGGGGAGTACAGAACTACCGGGGTTTGACGTATATCTGTGCCCTTA 6470

Query 414 ATATACAATTGGGCAGGAATGGCGTCACCCACCATAACCACGATTCCAAAAATGCTT 473
|
Sbjct 6469 ATATACAATTGGGCAGGAATGGCGTCACCCACCATAACCACGATTCCAAAAATGCTT 6410

Query 474 GCAACAGTTGTTCCCTACGTAAACAGAAACATACTGTCCCCATACACATCTGCAGAC 533
|
Sbjct 6409 GCAACAGTTGTTCCCTACGTAAACAGAAACATACTGTCCCCATACACATCTGCAGAC 6350

Query 534 ATTTGCAAAATATCAGGATATTTACAAATGGATTGACAAATATCTAAAGGCACCTCACTT 593
|
Sbjct 6349 ATTTGCAAAATATCAGGATATTTACAAATGGATTGACAAATATCTAAAGGCACCTCACTT 6290

Query 594 TTGGTTTCCTGCAATGCACCAAGTCCATAGCTCCATAGCCAGTATCAATCATATCACCA 653
|
Sbjct 6289 TTGGTTTCCTGCAATGCACCAAGTCCATAGCTCCATAGCCAGTATCAATCATATCACCA 6230

Query 654 TCCTCAATAGGGGTGTTTACTAGTTCCAAAGGAGGACAGTCCCCGTAGATACATTATTG 713
|
Sbjct 6229 TCCTCAATAGGGGTGTTTACTAGTTCCAAAGGAGGACAGTCCCCGTAGATACATTATTG 6170

Query 714 GGCTTGCAATGCCCTTCCCTTACCCAGTGTCTCCCAATGGCGGGAACACAGCCTATAATG 773
|
Sbjct 6169 GGCTTGCAATGCCCTTCCCTTACCCAGTGTCTCCCAATGGCGGGAACACAGCCTATAATG 6110

Query 774 CACAACGTGTCTGTATTATAATCCACAGACACATTATCCCTACTGTCTTATTGGTGGTT 833
|
Sbjct 6109 CACAACGTGTGTCTGTATTATAATCCACAGACACATTATCCCTACTGTCTTATTGGTGGTT 6050

Query 834 GATGAAATGGTGAGTTTTCAGTATCATCTGTCTATTATATAATGGGTGTCCACTAATA 893
|
Sbjct 6049 GATGAAATGGTGAGTTTTCAGTATCATCTGTCTATTATATAATGGGTGTCCACTAATA 5990

Query 894 CCAACACCCAATGGCTGGCCCTGC 912
|
Sbjct 5989 CCAACACCCAATGGCTGGCCCTGC 5965

```

Figure 4.10: BLASTN search result from a sequence generated by M13 reverse primer, showing 99% homology to HPV 39. Query sequence shows the sequence submitted for search in database while the subject is the identified organism, HPV 39.

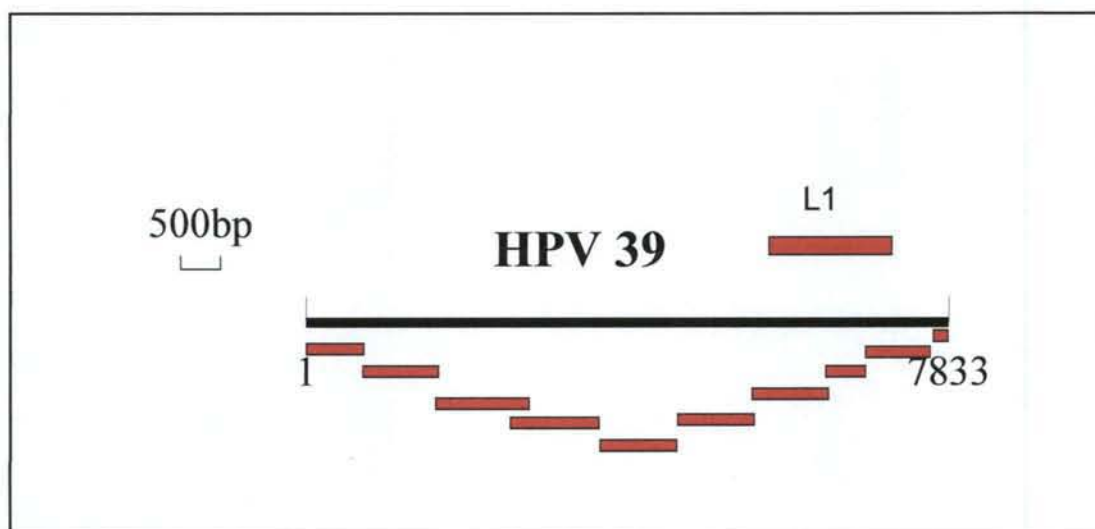


Figure 4.11: Partial sequences of HPV 39 generated by primer-walking. Sequencing was initiated by M13 primers in the L1 region at the restriction site of *EcoRI*.

*			
HPV39	AAATGCAACAGATACAGGTTTCAGACCTGGCAGACTTTATTGATGATTCCACAGATATTTGTGTACAGGCAGAGCGTGAGA	1120	
15.2	-----t-----	1120	
15.3	-----t-----	1120	
*			
HPV39	AGATGATGCAACCGGTACCTGCTGGTCATATTTTCGATAATTATATGAGAAATGCATTAGATGGGTATGCAATAAGTTTAG	2560	
15.2	-----a-----	2560	
15.3	-----	2560	
*			
HPV39	CCGCCGCCGTCTCGAAAGCGACCTCGACAGTGTGCAGTCACAGAGCCCACTGAGCCCGACGGAGTGTCCCTGGACCATCT	3600	
15.2	-----	3600	
15.3	-----t-----	3600	
*			
HPV39	TAACAACCCACTCCACAGTAACAGTACAGGCCACAACACAAGACGGTACCTCAGTTGTGGTAACACTACGCCTATAATAC	3680	
15.2	-----	3680	
15.3	-----t-----c-----	3680	
*			
HPV39	ATTTAAAAGGTGACAAAAATGGTTTAAAATGTTTAAAGATATAGACTACAAAAATATGACACATTGTTTGAAAATATTTCA	3760	
15.2	-----a-----a-----	3760	
15.3	-----a-----a-----	3760	
*			
HPV39	TGTACCTGGCATTGGATACGGGGTAAGGGAACCAAAAACGCTGGCATATTAAGTGTACATATGCCACAGAGTCACAACG	3840	
15.2	-----a-----	3840	
15.3	-----a-----	3840	
*			
HPV39	GGATATTGTGTATGTATATTGTATACATACTACATAGATGATATTATTGGTATTTTGGTGTGGTTTGGTGTGTGTATAT	4000	
15.2	-----g-----a-----	4000	
15.3	-----g-----a-----	4000	
*			
HPV39	AATGGATATGATATAGTACTGTATATGTATGTGCATTGTGCATAACTACTGTACATAGCTTTTTATATTTTTTTTGTGTTA	4240	
15.2	-----g-----	4240	
15.3	-----g-----	4240	
*			

Figure 4.12: Alignment of HPV 39 (GenBank accession number M62849) with the HPV 39 sequences from clones, 15.2 and 15.3. Point mutations were detected throughout the genome. The mutations are highlighted in yellow. The asteriks (*) inbetween sequences indicate missing regions where no mutations were identified and therefore the sequence alignments for those regions are not shown.

4.3 Discussion

The detection of HPV types in clinical samples is dependent on the methodology applied. The Roche Linear Array HPV genotyping test used in Chapter two detected 12 different HPV types in the specimen HH015 (Figure 4.5). The pyrosequencing method used in this chapter identified only seven HPV types from the same specimen, namely HPV 16, 30, 39, 40, 56, 74 and 86. The differences in the HPV types identified by the two detection methods are highlighted in Table 4.2. Among the seven HPV types detected by pyrosequencing analysis, HPV 16, 39, 40 and 56, are included in the Linear Array detection test. However, HPV 56 was not detected by the Roche Linear Array HPV Genotyping test although it was part of the panel of HPV types that could be detected. It is possible that HPV 56 was not detected by the PCR-based Roche Linear Array HPV Genotyping test due to low levels. PCR-based methods may not be effective for HPV types with low DNA levels, especially in HH015 sample which had multiple HPV infections. Primer competition between different genotypes in PCR reactions could also explain the failure to detect HPV 56 by the Roche Linear Array HPV Genotyping test. This implies that PCR does not equally amplify the different HPV genotypes. According to the Roche Linear Array Genotyping test's manual, the test can detect HPV DNA at various concentrations but detection is also dependent on the HPV type. Therefore, assuming that HPV 56 had a low viral load, it was not amplified to detectable levels. On the contrary, pyrosequencing is more sensitive (Wicker *et al.*, 2006, Gharizadeh *et al.*, 2005), therefore may be more reliable for samples negative with PCR methods. This is because in pyrosequencing, a single ssDNA fragment was clonally amplified to generate several copies of the same DNA molecule. This occurred for every fragment of DNA generated during the nebulisation process (Figure 4.2) thereby ensuring that each strand was sequenced. Unlike in PCR reactions where nucleotides may get depleted after a number of reactions, during pyrosequencing, the nucleotides are constantly being added at the start of each cycle. This ensures the continuity of sequencing process without the limitations of reagents. The RCA on DNA before pyrosequencing, which was not performed before PCR, improved the DNA concentration in sample hence this could have improved HPV 56 DNA level. This may have contributed to the detection of HPV 56 by pyrosequencing and not by Roche Linear Array HPV Genotyping test.

In addition to HPV 56, three types (HPV 30, 74 and 86) not included in the Roche Linear Array HPV Genotyping test were detected. The viral load of these types was unknown but one can speculate their viral load may have been present at intermediate or high level for them to be identified. Although they were detected, the coverage of the sequences from these genomes was low compared to that of HPV 39 (Figure 4.7). This may imply that the DNA concentration for these three HPV types was lower than that of HPV 39. The question of sensitivity of pyrosequencing with regard to viral load could be solved by quantitating each HPV type DNA level and compare to the results obtained by pyrosequencing. However, this was not achieved due to time constraints.

Pyrosequencing failed to detect nine types (HPV 45, 52, 53, 55, 59, 70, 71, 81 and 84) in Table 4.2 that were detected by the Roche Linear Array HPV Genotyping test. Although the pyrosequencing was able to detect multiple infections in the specimen, some of the high-risk HPV types known to be present were not detected. While it is possible that the viral load of each type may have influenced its successful identification, there is the possibility that pyrosequencing may not effectively identify all types in multiple infections. This may be due to differential viral loads of the different HPV genotypes. The detection of multiple genotypes in a single specimen may be a challenge for the pyrosequencing of whole genomes as opposed to PCR amplicon sequencing. In previous studies on HPV genotyping by pyrosequencing, HPV PCR products amplified by general primers GP5+/6+ and MY09/11 were used (Gharizadeh *et al.*, 2001). The PCR products were primed with the GP5+ primer before pyrosequencing. This resulted in accurate genotyping of samples. PCR products are easier to sequence because of small size. The use of HPV general primers improved sequencing by increasing the sensitivity for HPV DNA. However, in the presence of non-specific amplification in some of the samples or multiple types of HPV, pyrosequencing resulted in mixed signals (Gharizadeh *et al.*, 2001). This was improved by using HPV-specific primers in pyrosequencing (Gharizadeh *et al.*, 2005). In comparison, the pyrosequencing in our study did not use HPV-specific primers and complete HPV genomes were sequenced after non-specific RCA of circular DNA from specimen. Therefore, this could have contributed to identification of only a few types (seven) compared to the 12 detected by the Roche Linear Array HPV Genotyping test. Samples with multiple infections may be particularly problematic due to the multiple sequence signals from all the genotypes in one specimen. This could have resulted in mixed sequence peaks which were difficult to analyse on the system

(Ahmadian *et al.*, 2006). Multiple signals could have been generated in the sequencing of the conserved regions like the E1, E2 and L1, where the simultaneous addition of identical nucleotides was possible. This could have made it difficult to determine the correct number of incorporated nucleotides (Ronaghi, 2001). According to Gharizadeh and colleagues (2001), genotyping by pyrosequencing was found to be most effective in single HPV infections although typing was also possible when there were few dominant types and low background from other infecting HPV types (Gharizadeh *et al.*, 2001). This could explain why only a few HPV types were detected by pyrosequencing analysis, which seemed to reflect the HPV types that seemingly, had higher viral loads compared to the others.

The HPV 39 band was the most intense of all the HPV types detected by the PCR-based Roche Linear Array HPV Genotyping test. The higher intensity of the HPV 39 band in the Roche Linear Array HPV Genotyping test (Figure 4.5) indicated a higher viral load in comparison to the other HPV types present in the specimen. Results from pyrosequencing showed a complete genome of HPV 39 (Figure 4.7d). The presence of HPV 39 in abundance could have contributed to its full length sequencing compared to the short sequences from other HPV types (Figure 4.7b, c, f and g). HPV 40 genome was almost fully sequenced as there were only a few (73) bases missing in the assembled sequence (Table 4.1 and Figure 4.7e). HPV 16 was completely sequenced but assembled into several overlapping fragments that constitute a complete genome (Figure 4.7a). The other HPV types (74, 86, 30, 56) were only partially sequenced (Figure 4.7b, c, f and g, respectively).

Several of the sequences from pyrosequencing were identified as human DNA. This was expected since the origin of sample was human cells. As for supposed hits to other organisms, with the exception of bacteria and endogenous viruses, these represented similarities with human DNA rather than being present in specimen DNA. *Bacteroides fragilis* is a commensal organism within the human gut, which causes opportunistic infection in HIV immunosuppressed individuals (Wexler, 2007). Opportunistic infections by this organism can occur at any site, including the genital area (Wexler, 2007). Therefore the presence of bacteria could have been from the sampling. *Prevotella intermedia* is a gram negative anaerobe which can cause infections of the oral and genital tract (Finegold, 1995). The organism was therefore presumed to be present in the genital

tract of the individual. HERV-K is an endogenous virus that is commonly found in human DNA and does not cause infection (Boller *et al.*, 1993). The other organisms which were identified (Figure 4.6) from other sequences indicated sequence similarities to DNA rather than their presence in human sample. The *Pan troglodytes*, for example, have a common ancestry with human hence the homology to some of the DNA sequences.

The abundance of HPV 39 in the sample provided competitive advantage over other types digested with *EcoRI* when cloning was performed; therefore the RCA clones sampled had HPV 39 sequence. This was also reflected in the DNA bands shown in agarose gel after restriction enzyme digestion (Figure 4.8). The more abundant HPV 39 'suppressed' the presence of other HPV types at low viral load. As a result, HPV 39 was the only type investigated for recombination within HPV genomes. There were no recombination events detected within HPV 39 sequences analysed. These observations were from only two clones and therefore may not reflect the correct impression. More cloned HPV genomes need to be analysed. Due to time constraints, analysis was performed on DNA cleaved by one restriction enzyme, *EcoRI*.

Although no recombination was detected in HPV 39 sequences, several point mutations were identified within different genomic regions. The two sequences analysed revealed the occurrence of point mutations (Figure 4.12). This confirms other observations that sequence variations are common within HPV genomes. In a study of variation among HPV types by Stewart and colleagues (1996), HPV 39 sequences showed two nucleotide changes within the L1 ORF (Stewart *et al.*, 1996). In our study, several point mutations were noted within various ORFs including E1, E2/E4, E5, and L2. These sequence variations potentially signified variants of HPV 39, which increase genomic diversity of HPV. In the event of recombination, some regions which belong to other HPV sequences would have been identified within the HPV 39 genome. For example, a 20bp region from HPV 16 identified within HPV 39 genome sequence could have been an indication of recombination between HPV 16 and HPV 39.

Most HPV infections appear to be clonal (Levi *et al.*, 2002; Vinokurova *et al.*, 2005), but co-infection with different HPV types is very common only in HIV infected individuals (Moodley *et al.*, 2006). Therefore, there is a unique opportunity for new recombinant

HPV types to be generated within the host (Angulo and Carvajal-Rodríguez, 2007). When considering the high prevalence of HIV globally, HPV recombination might be favoured due to multiple persistent HPV infections in HIV positive women (Varsani *et al.*, 2006).

The discovery of the BPCV1 was considered to be a result of possible recombination between PVs and PyVs (Woolford *et al.*, 2007). The detection of possible recombination between these two subfamilies of *Papovaviridae* may signify that recombination is more likely to occur between HPVs, which are more closely related to each other than PVs are to the PyVs.

Since the methodology for detecting recombination has been identified, it is promising that recombination can be detected among individuals with multiple HPV infections and HIV immunosuppression.

The presence of additional HPV types as detected by pyrosequencing indicates the need to incorporate additional detection methods in HPV screening. It also suggests increased diversity of HPV types infecting women with HIV co-infection. The current methods of choice are HC2 assay and the Roche Linear Array HPV Genotyping test. Although these methods detect most clinically relevant HPV types, some types can easily be missed due to non-inclusion in the test method. The HPV types detected by the pyrosequencing method could be incorporated in the current Roche Linear Array HPV Genotyping test to increase the spectrum of HPVs detected in HIV infected individuals. Since HPV 30, 74 and 86 were detected in one specimen, their prevalence in this HIV population required investigation by an appropriate method. This was followed on by the application of type-specific PCR on all the specimens, refer to Chapter 5.

Chapter 5

PCR screening of clinical specimens for additional HPV types

5	INTRODUCTION.....	90
5.1	MATERIALS AND METHODS.....	91
5.1.1	<i>Primer design.....</i>	91
5.1.2	<i>Type-specific PCR amplification of HPV DNA.....</i>	91
5.1.3	<i>Cloning and sequencing.....</i>	92
5.2	RESULTS.....	93
5.2.1	<i>Type-specific PCR amplification.....</i>	93
5.2.2	<i>Prevalence of additional types (HPV 30, 74 and 86).....</i>	95
5.3	DISCUSSION.....	96

5 Introduction

The use of PCR-based methods in HPV detection has had huge successes in clinical assessment of HPV infection cases. The Roche Linear Array HPV Genotyping test is one of these methods; however, despite its successful application in HPV detection, it has its limitations in identifying all types in multiple infections. The inability of this method to distinguish HPV 52 from the other closely related HPV types, for example, also limits accuracy in reporting HPV 52 prevalence.

The application of HPV type-specific PCR in diagnostics has proved to be very useful in improving the accuracy of results obtained in epidemiological studies. This method is often used alongside other detection methods and is applied to confirm or refute certain results due to its increased sensitivity. There is no competition for primers as observed in most PCR reactions using multiple primers and it is appropriate for amplification from samples with low HPV DNA viral load. However, one of the drawbacks for its application is that the primers for specific genotypes are 'in-house' developed. There are no commercially available type-specific primers for each genotype. Another disadvantage for using type-specific primers is the need for multiple parallel PCR reactions for amplification of HPV types and post-amplification identification of types. This makes this process labour intensive. For these reasons, type-specific PCR can not be applied to huge samples in epidemiology, diagnostic and clinical studies (van den Brule *et al.*, 2002; Lin *et al.*, 2008).

In this chapter, a study of the prevalence of three additional HPV types (HPV 30, 74 and 86) is reported. These types are not included in the cocktail of probes detected by the Roche Linear Array HPV Genotyping test, but were detected by pyrosequencing of DNA from one clinical specimen, HH015 (section 4.2.2). The prevalence of these types among all the women in the study was not known. Therefore, it was relevant to investigate prevalence of HPV 30, 74 and 86 by type-specific PCR among all the women.

5.1 Materials and Methods

5.1.1 Primer design

Primers were designed to target only the HPV type under investigation. Therefore primers specific for HPV 30 (GenBank accession number X74474) were designed to target a short amplicon of 345bp within the E6 gene. Primers that amplified HPV 74 subtype AE10 (GenBank accession number AF436130) and HPV 74 (GenBank accession number U40822) were designed to amplify a 329bp (Table 5.1). HPV 86F and HPV 86R primers were designed from a 241bases sequence obtained from pyrosequencing in the previous chapter. The primers amplified a 236bp fragment and the primer sequences corresponded to HPV 86 (GenBank accession number AF349909) at positions within the E6 gene. The primers are shown in Table 5.1.

5.1.2 Type-specific PCR amplification of HPV DNA

The PCR mix for amplification of HPV 30 was set up as follows: 1X GoTaq DNA polymerase buffer (Promega, Madison, USA), 1mM MgCl₂, 200μM dNTPs, 0.65U GoTaq DNA polymerase enzyme (Promega, Madison, USA), 2μl of DNA template and 400μM of each primer in a 25μl reaction volume. DNA template used was extracted from each of the 109 cervical specimens. The dNTP solution contained nucleotides (dATP, dCTP, dGTP and dTTP) for incorporation during primer extension. The MgCl₂ concentration was adjusted to 2mM for the amplification of HPV 74. The enzyme catalysed the addition of nucleotide bases at the 3'end of the DNA strand. PCR cycling conditions for amplification of HPV 30 and 74 were set as follows: 94°C for 10 minutes for denaturation of ds DNA, 35 cycles consisting of 94°C for 40 seconds, 55°C for 40 seconds for primer binding to DNA template, 72°C for 40 seconds for extension of bound primer and a final hold at 72°C for 10 minutes for the final elongation of bound primer. PCR mix for amplification of HPV 86 was the same as for HPV 30 amplification, except for the difference in primers. The annealing temperature was adjusted to 53°C for the amplification of HPV 86 while the cycling conditions were kept the same as above. The amplified products were electrophoresed on a 2% agarose gel as outlined in Appendix Aiii.

Table 5.1: Type-specific primers for the amplification of HPV 30, 74 and 86

Primer	Sequence	Position	Amplicon size (bp)
HPV30F	5'-TGAGGTACAAGAAACATCGTTGC-3' (23)	158-180 ^c	345
HPV30R	5'-CGTACGTGATATTCTGTGAAACC-3' (23)	503-481 ^c	
HPV74F	5'-TGCTGGACAACATGCATGGAAAA ATCCTAC-3' (31)	418-448 ^a , 3365-3395 ^b	329
HPV74R	5'- CCTCTGTACCTGTATTTCCGCCATGT -3' (27)	745-719 ^a , 3696-3672 ^b	
HPV86F	5'-GTTTGTAGGAGTGTGGCATCC-3' (21)	29-10 ^c	236
HPV86R	5'-TTGTACTGCCAAGTTTCTGTCC-3' (22)	7781-7802 ^c	

^a Position of primers in the HPV 74 AE10 subtype (AF436130) within the E6/E7 genes. ^b Position of primers in the HPV 74 (U40822) within the L1 gene. ^c Primers located within the E6 gene.

5.1.3 Cloning and sequencing

The cloning and subsequent sequencing was performed to verify that the correct amplicon had been amplified by the respective primers. PCR amplicons from HPV 30, 74 and 86 (amplified from specimen HH015) were excised from gel and purified according to protocol in Appendix Av. The purified amplicons were ligated into pGEM[®]-T Easy vector according to procedure in Appendix Avi using the pGEM[®]-T Easy cloning system (Promega, Madison, USA). The plasmid vector was transformed into *E.coli* (Appendix Avii) for replication and DNA was isolated from 5 ml bacterial culture (Appendix Aii). Purified DNA was sequenced using the M13 universal primers (Appendix Bii) to confirm that the correct HPV type had been amplified from specimen. Sequencing was performed at Macrogen, Korea, according to the chain termination method as described in Appendix Aviii. Analysis of sequences was done using Chromas version 2.01, 2001 and DNAMAN program for windows version 5.2.9 (Lynnon Biosoft, © 1994-1999).

5.2 Results

5.2.1 Type-specific PCR amplification

PCR screening of all the specimens was performed using type-specific primers in order to determine the prevalence of each of HPV 30, 74 and 86 among 109 HIV positive women. Primers specific to each of the HPV types 30, 74 and 86 were used to amplify short fragments from clinical specimens from 109 HIV positive women. Amplicons from specimen HH015 which were cloned and sequenced showed that correct HPV types were amplified (Table 5.2). For each sequence searched by BLASTN (www.ncbi.nlm.nih.gov/blast), other related HPV types were also identified, although with less homology (Table 5.2). A phylogenetic analysis of these HPV types showed that they belong to the same species as shown in Figure 5.4.

Type-specific primers used in the amplification of HPV 30, 74 and 86 successfully amplified the expected PCR products from other specimens positive for each respective HPV type. Figures 5.1-5.3 illustrate the results of the amplification.

Table 5.2: Blast results of clones of PCR amplicons from different HPV types.

Amplicon from	length of sequence (bp)	% identity	Identity to any other HPV type
HPV 30	400	98	88%-HPV53
HPV 74	369	96	90%-HPV 74cds
HPV 86	238	96	79%-HPV 84

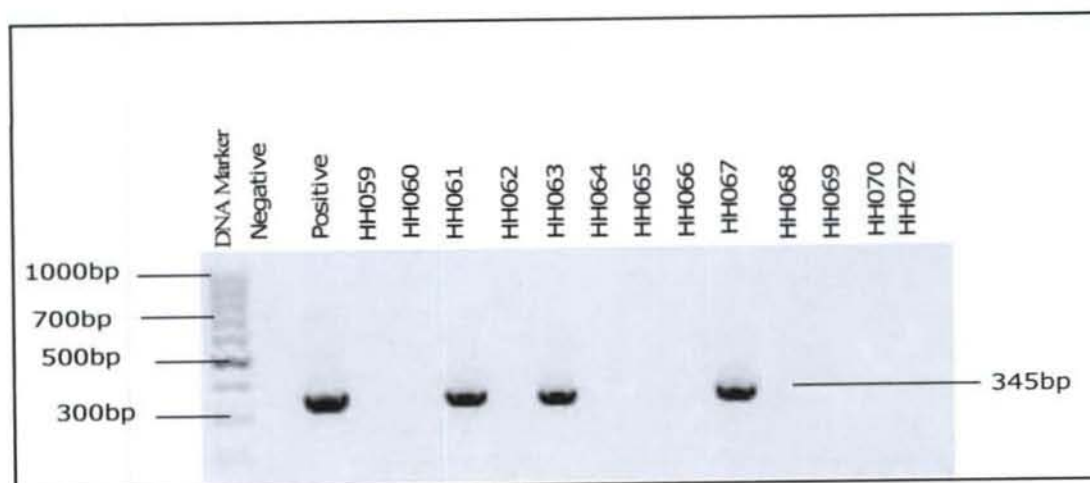


Figure 5.1: Representative agarose gel analysis after PCR amplification of HPV 30 from clinical specimens. A negative control with water and a positive control using DNA from HH015 were included in each amplification. The specimens HH061, HH063 and HH067 were positive for HPV 30.

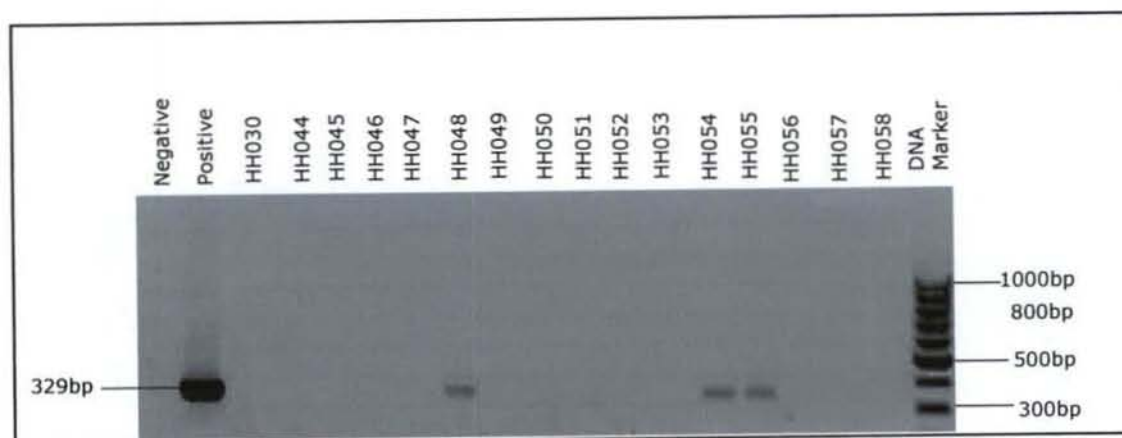


Figure 5.2: Representative agarose gel of the amplification of HPV 74 from 109 clinical specimens. The three specimens in this amplification that were positive for HPV 74 were HH048, HH054 and HH055.

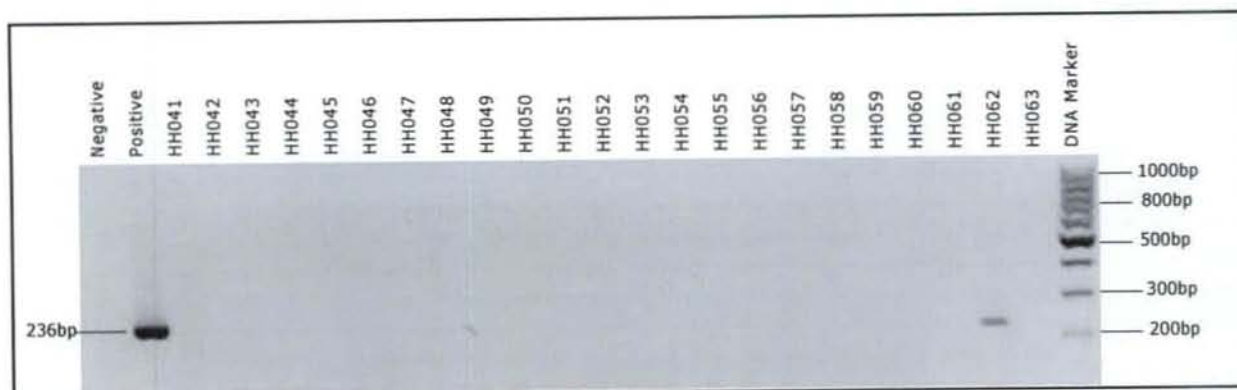


Figure 5.3: Representative agarose gel after the amplification of HPV 86 from clinical specimens. Specimen HH062 was positive for HPV86 in this figure.

5.2.2 Prevalence of additional types (HPV 30, 74 and 86)

Each of the 109 specimens was screened for the presence of each of the three HPV types. The specimens positive for each of the three HPV types are shown in Table 5.2. HPV 30 had the highest prevalence (15%, 16/109), HPV 74 prevalence was 13% (14/109) and 5% (5/109) of the women were infected with HPV 86 (Table 5.3).

Table 5.3: The distribution of HPV 30, 74 and 86 among 109 HIV positive women

Description	Patients	Total	
		n=109	%
HPV negative		81	74.3
HPV 30	hh015, hh022, hh031, hh034, hh040, hh041, hh051, hh054, hh055, hh061, hh063, hh067, hh076, hh082, hh095, hh102	16	14.7
HPV 74	hh015, hh020, hh022, hh034, hh048, hh054, hh055, hh069, hh074, hh083, hh087, hh097, hh107, hh111	14	12.8
HPV 86	hh011, hh015, hh034, hh062, hh099	5	4.6

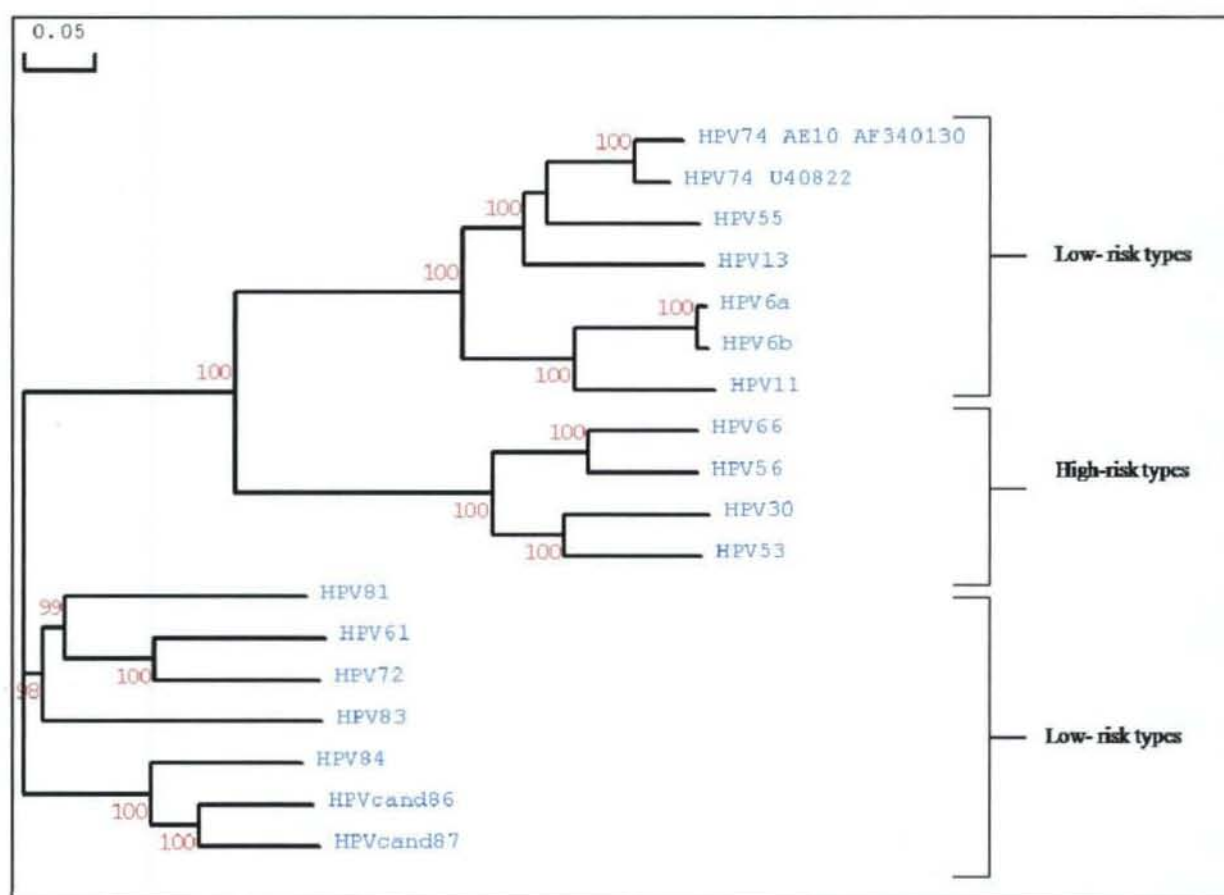


Figure 5.4: Phylogenetic tree of different HPV types. The full length genome sequences obtained from GenBank (NCBI) were aligned using the Jukes and Cantor joining and a bootstrapping of 100. HPV 30 is closely related to the high-risk HPV types 53, 56 and 66, a high-risk clade. HPV 74 is related to the low-risk types 55 and 13, while HPV 86 shows close relationship to HPV 87 and 84. Both HPV 74 and HPV 86 are phylogenetically classified as low-risk HPV types.

5.3 Discussion

In Chapter 4, three HPV types (30, 74 and 86) were detected in one specimen using pyrosequencing technique. These were not detected by the Roche Linear Array HPV genotyping test. Therefore it was relevant to investigate the prevalence of these HPV types in all the other women. To achieve this, type-specific PCR was used to amplify each of the three HPV types.

HPV 30 had the highest prevalence (15%, 16/109) among the three HPV types (Table 5.3). This is significant, as its prevalence was higher than the prevalence of HPV 16 (14%, 15/109) in this population study. This may indicate that HPV 30 could be a commonly occurring type among HIV positive women. It was first isolated from laryngeal carcinomas and also detected in a few genital lesions (Kahn *et al.*, 1986; IARC, 2007). This type has not been reported in many epidemiological studies due to its exclusion in the commercial kits for HPV detection. HPV 30 is not included in these kits because its risk for cervical cancer has not been proven since it has been found mostly in benign lesions. Two studies have reported on the prevalence of HPV 30 and only one of them was on HIV positive women (Banura *et al.*, 2008; Singh *et al.*, 2009). Among the HIV positive women, overall HPV 30 prevalence was 6% (40/647) and in women with HSIL, it was 9% (6/69) (Singh *et al.*, 2009). The prevalence in the general population was reported as 9% (9/95) (Banura *et al.*, 2008).

HPV 30 is classified in the genus alpha-papillomavirus, species six, together with HPV 53, 56 and 66. Among the species six group, HPV 56 is the only type categorised among the high-risk while 53 and 66 are probable high-risk types (Schiffman *et al.*, 2009) (Chapter 1, section 1.5.1). A phylogenetic analysis of HPV 30 indicated that it is closely related to HPV53 (Figure 5.4) and therefore should probably be considered as a high-risk type. Results from Singh and colleagues' (2009) study indicated that infection with HPV 30 among HIV positive women led to HSIL (Singh *et al.*, 2009). For this reason, its prevalence in the population should be monitored. This can be achieved when HPV 30 has been included in the commercially available HPV detection kits that may be used in cervical screening (HC2 and Roche Linear Array HPV Genotyping test).

HPV 74 was slightly less common than HPV 30, with a prevalence of 13% (14/109). It is a mucosal HPV related to the low-risk types in species ten of the genus alpha-papillomavirus; HPV 6, 11, 13, 44 and 55 (IARC, 2007; Longuet *et al.*, 1996). A phylogenetic tree of these types indicated that HPV 74 is more closely related to HPV 44 and 55 (Figure 5.4). This HPV type is a non-oncogenic mucosal type associated with LSIL and has been reported in vaginal intraepithelial neoplasia (VaIN) (Wheeler, 2002; Safaeian *et al.*, 2007). HPV 74 was initially isolated from immunosuppressed women who had renal transplants and had LSIL of the vagina. It was not isolated from any immunocompetent patients (Longuet *et al.*, 1996). Recently, however, HPV 74 has been reported in 2/82 (2%) of HIV positive women and a lower prevalence (1%, 8/868) in HIV negative women (Banura *et al.*, 2008). This may reflect that HPV 74 occurs infrequently and can effectively be cleared by the host immune system of immunocompetent individuals (Longuet *et al.*, 1996). The presence of HPV 74 in renal transplant patients and HIV positive women indicate the susceptibility of immunocompromised women to HPV 74. Therefore, HPV detection methods may need to include HPV 74 especially when monitoring HPV infection in HIV positive populations.

HPV 86 was the least common of the three types, with a recorded prevalence rate of 5% (5/109) among the HIV infected women. It is a mucosal low-risk HPV type responsible for genital warts as well as CIN (Wheeler, 2002). HPV 86 is classified into the genus alpha-papillomavirus, species three, and is phylogenetically related to the HPV 84, 87, 61, 72, 81 and 83 (IARC, 2007). Figure 5.4 shows the relationship between HPV types in species three of the alpha-papillomavirus genus. In a recent study among HIV positive women, the overall HPV 86 prevalence was 35/647 (5%) and in women with cervical lesions, was 14/212 (7%). Of the 69 women with HSIL, six of them (9%) were positive for HPV 86, thus linking HPV 86 to HSIL in HIV positive women (Singh *et al.*, 2009). The low prevalence of HPV 86 in Singh and co-workers' (2009) study and our study may suggest that it is a rare type in HIV positive women and probably more rare in the HIV negative population.

The identification of these mucosal HPV types in HIV positive women may suggest that HPV diversity among these women is greater than anticipated. There could be more HPV types that are not necessarily associated with risk of cervical disease/cancer in immunocompetent women, but are relevant to immunosuppressed patients, that remain

undetected. Since type-specific PCR may not be practical to use in large epidemiological studies, current detection methods need to cover a wider spectrum of HPV types. Most detection methods include the high- and probable high-risk types but few of the low-risk types. More coverage of the low-risk HPV types and some types considered to be of undetermined risk (Table 1.3) in the detection methods would benefit screening for cervical disease among the HIV positive women.

Chapter 6

6 Conclusion

HIV positive women are infected with more HPV types than their HIV negative counterparts. The wider HPV range in these women demonstrates the need to continue the development of polyvalent vaccines for the prevention and treatment of cervical disease in women. Vaccination before sexual debut has been instrumental in the prevention of infection with HPV 6, 11, 16 and 18. However, the vaccines may not be effective against infection with other HPV types in HIV positive women who are at higher risk of infection than HIV negative women. Development of polyvalent prophylactic vaccines incorporating the most prevalent HPV types among HIV positive women would reduce the burden of cervical disease in South Africa. Ultimately, this would also decrease mortality from cervical cancer in the general population. The most frequently occurring HPV types, in this study in HIV positive women were HPV 61, 66, 53 and 58, which could be added to the HPV types targeted for vaccine development, especially for high HIV prevalence areas. The high prevalence of HPV types that are not included in the current vaccines emphasises the need to develop vaccines that protect against a wider range of HPV types. From this study, low-risk HPV 61 had the highest prevalence (24%, 26/109), indicating that it is a potential threat to the immunosuppressed women. The prevalence of the high-risk HPV 58 (17%, 19/109) and probable high-risk types 53 (17%, 18/109) and 66 (18%, 20/109) also potentially increases the vulnerability of these women to cervical disease/cancer. In addition to these, HPV 45, which was also found to be common among HIV positive women in a different study in Cape Town (Marais *et al.*, 2008), can be considered as a vaccine candidate in order to reduce HPV infection among HIV women.

The Roche Linear Array HPV Genotyping test is currently not equipped to identify all the HPV types in the HIV positive immunocompromised population. This was highlighted by the identification of HPV types not included in the Roche Linear Array HPV Genotyping test by using type-specific PCR (Table 4.2). As observed in this study, there were low-risk HPV types detected frequently among these women, therefore it would be beneficial

to include more HPV types in detection kits than are currently incorporated. The application of other detection methods in addition to the PCR-based method identified more HPV types, thus re-affirming the need for more HPV types to be included in diagnostic detection kits. The identification of HPV 30, a type closely related to the high-risk HPV 56, indicated the possibility that HIV infected women may have more 'unclassified' HPVs which can cause HSIL than those currently classified as high-risk types. Although HPV 30 is not among the high-risk HPV types, it can lead to HSIL in HIV positive individuals as was observed by Singh and co-workers (Singh *et al.*, 2009). The number of studies on HPV 30 and HSIL cases are however limited, therefore more studies on this type are required to determine its risk in cervical disease or cancer.

While PCR-based methods are considered sensitive, it is worth noting that different PCR methods can detect certain HPV types better than others, especially in multiple infections. For example, the Roche Linear Array HPV Genotyping test for detection of HPV 52 and 31 was found to be poor compared to the detection of these types using the LiPA detection method (Castle *et al.*, 2008). As a result, some HPV types may be underestimated and this certainly impacts on the prevalence of HPV types reported in epidemiological studies. The use of at least two HPV detection methods would be ideal to accurately report on HPV prevalence.

The use of the Roche Linear Array HPV Genotyping test alone may not be reliable for the estimation of HPV 52 prevalence. In research work, it would be beneficial to use another method for the detection of HPV 52 in samples positive for the mixed group, for example type-specific PCR. This method is very useful in confirming disputed results or can be used as a third method where two other methods are discordant. For the purpose of clinical management of women, it would be safer to assume presence of HPV 52 whenever a sample is positive for the mixed group (Castle *et al.*, 2008). The use of the recently developed BSGP5+/6+ primers can also improve the identification of HPV 52. In addition, these primers can detect HPV 30, 74 and 86 (Schmitt *et al.*, 2008). The coupling of the Roche Linear Array HPV Genotyping test genotyping with the BSGP5+/6+ primers could significantly increase the spectrum of HPV types detected. Therefore, this would significantly improve HPV monitoring among HIV positive women.

Pyrosequencing of DNA from HPV and HIV infected individual indicated the presence of more HPV types in addition to those detected by Roche Linear Array HPV Genotyping test. Despite its inability to detect all the types that were genotyped by the Roche Linear Array HPV Genotyping test, pyrosequencing can effectively be used in identification of novel HPV types as well as those currently not included in PCR-based methods. It can also give an idea of differential viral loads. The efficacy of this method for detecting multiple HPV infections may be improved by the use of multiple HPV-specific primers during sequencing (Ahmadian *et al.*, 2006). Pyrosequencing can also be applied to the identification of HPV 16 variants. This would benefit infected women since presence of HPV 16 variants can be used as a marker for cervical disease among HIV positive women. Successful identification of women at risk would therefore allow necessary intervention to prevent cervical cancer.

Although we did not detect recombination in the analysed HPV genome, there is a possibility of recombination occurring in HIV positive women with multiple HPV infections. Through the application of RCA and metagenomic analysis, which includes pyrosequencing, it is possible to analyse several HPV genomes for recombination. The analysis of more HPV genomes can increase the possibility of detecting recombination. In our study, however, we were able to analyse only one HPV genome.

The genomic diversity of HPV can be fully understood only when efficient diagnostic tools are available. Identification of HPV 16 variants by sequencing the LCR indicated the high level of sequence variation, which could give rise to even more diverse HPV genomes. This high level of sequence variation within the LCR was expected since the LCR is under less mutation restriction when compared to the coding regions. The observation of new mutations in this region suggests that frequent point mutations occur within the LCR. The mutations observed in HPV 16 may also suggest that a similar pattern of variation may occur in other HPV types. An investigation into the diversity of other HPV types that are phylogenetically unrelated to HPV 16 showed presence of limited genomic diversity. This was observed in variants of HPV 53, 56 and 66, which have also been identified among South African women (Prado *et al.*, 2005). The presence of point mutations within HPV 39 genome indicated sequence variations that can increase genomic diversity among HPV types.

In conclusion, HIV positive women have increased diversity of HPV infection from both low- and high-risk HPV types. This necessitates constant monitoring of these women to prevent progression to cervical cancer. All the known variants of HPV 16 (E, A, AA, and Af) were identified among this group of women. There were no recombination events observed in the analysed HPV 39 genome. There is need for analysis of more HPV genomes and different HPV types for identification of recombination in HPVs.

Appendix A

Standard Methods

i. Bacterial growth media and growth conditions

Bacterial cultures were grown in Luria-Bertani (LB) broth containing 100µg/ml of the antibiotic ampicillin. The ampicillin was added for selective growth of transformed cells. The LB consisted of the following: 5g bacto-typtone, 5g yeast extract, 10g NaCl and made up to 1L with distilled water and autoclaved. Luria-Bertani agar (LA) was also made up with the same components, and in addition contained 15g/L of agar. All growth was incubated at 37°C for 18 hours, with liquid cultures incubated on a shaker. The bacterial strains used for transformation were DH5α strain of *E.coli* (Invitrogen, USA).

ii. Small scale DNA extraction

DNA was isolated from 5ml overnight (18 hours) bacterial cultures according to QIAprep Spin miniprep Kit protocol (Qiagen, Gaithersburg, USA). This extraction method is based on the alkaline hydrolysis of cells in Sambrook *et al.*, 1989. After centrifugation, the pelleted cells were resuspended in buffer containing Tris-Cl, EDTA RNase A. The cells were lysed by incubation in alkaline lysis buffer with 200mM NaOH and 1% SDS to release proteins and nucleic acids. The solution was neutralised with potassium acetate before centrifuging. The RNase removed any contaminating RNA and SDS denatured proteins, which were removed in a pellet by centrifuging. In the presence of sodium salts, the DNA in solution was bound to a column of silica membrane and precipitated with isopropanol. Any contaminants were washed through the membrane. Precipitated DNA was washed with buffer containing ethanol and eluted with Tris-Cl-EDTA (TE) buffer by releasing bound DNA into solution.

iii. Agarose gel electrophoresis of DNA

Gel electrophoresis of DNA was performed on 0.8% agarose gels at 100 volts for genomic DNA and fragments larger than one kilobase (kb). DNA separates on agarose gel according to size of fragment. Electrophoresis of PCR amplicons of 500bp or less was carried out on 2% agarose gels at 120 volts. The separation of larger fragments (1kb or larger) was performed using 0.8% agarose gel. All agarose gel contained ethidium

bromide at a concentration of 0.5µg/ml for DNA visualization on a UV transilluminator at 260nm (Sygene). 1X Tris-Borate EDTA (TBE) was the buffer medium of electrophoresis. DNA fragment size was estimated by using the one kb O'Gene Ruler (range 10000bp-250bp) and 100bp O'Gene Ruler (range 1000bp-100bp) (Fermentas, USA) for large and small DNA fragments, respectively.

iv. Restriction Enzyme Analysis

Genomic and plasmid DNA was digested with *EcoRI* restriction enzyme (Fermentas, USA) in a 20µl reaction mix consisting of 2µl (1X) *EcoRI* buffer (Fermentas, USA), 0.5µl *EcoRI* enzyme, approximately 300ng DNA and made up to 20µl with sterile distilled water. The mixture was incubated in 37°C water bath for four hours. The enzyme was heat inactivated at 72°C for 5 minutes after the reaction. Digested DNA was electrophoresed on the appropriate percentage gel agarose (see iii). Other enzymes used for digestion of genomic DNA were *BamHI* and *BglI*.

v. Gel purification of DNA

DNA was purified from agarose gel slices according to the protocol of the Gel Purification Kit (Qiagen, Gaithersburg, USA). After restriction enzyme digestion or PCR, DNA was loaded onto the relevant agarose gel and electrophoresed, visualized on the long wavelength (345nm) and excised using a sterile blade. The excised gel slice was put into a 2.0ml eppendorf tube and purification was performed as outlined in above named kit. The gel slice was dissolved at 50°C in buffer at pH 7.5. The DNA was bound to a column of silica membrane and washed with ethanol for precipitation and eluted in TE buffer. DNA was either sequenced or used in subsequent reactions.

vi. Ligation reactions

Gel purified DNA was used in all ligation reactions in a 10 µl reaction volume consisting of vector to insert ration of 1:3 for small DNA fragments and 1:20 for large (>3kb) DNA fragments. The vector used for cloning PCR products was pGEM®-T Easy (Promega, Madison, USA), and for cloning of whole genome of HPV, pUC19 (Amersham Biosciences) was used. pUC19 was linearised with *EcoRI* (see iv) and dephosphorylated at ends to prevent self-ligation. DNA was ligated using 1U DNA Ligase enzyme (Promega, Madison, USA) in the presence of 1X Ligase buffer (Promega, Madison,

USA). The ligation reaction was carried out at 4°C for 18 hours. DNA ligase enzyme catalysed the formation of phosphodiester bonds between two DNA strands.

vii. Transformation

Competent *E.coli* DH5α cells (Invitrogen, USA) were used for all transformation experiments. Transformation was carried out by using 20µl of competent bacterial cells with 5µl of ligation mix in a 1.5ml eppendorf tube. The tube was chilled on ice for 15 minutes before heat-shocking at 37°C in a water bath for 1min and chilled on ice again for two minutes. 500µl of LB were then added to the transformation mix and incubated for a further 1 hour at 37°C, after which 100µl of the mix were plated out onto a LA plate with Ampicillin as a selective marker, 50 µl of 0.1M IPTG and 10 µl of 2% X-gal. Incubation of LA plates was as mentioned in (i) above.

viii. Sequencing of clones and sequence analysis

Purified miniprep DNA was used for sequencing using the Sanger chain termination method, at Macrogen, Korea. The sequencing technique is based on the dideoxy chain termination method. The sequencing method required a template DNA to be sequenced, primers to initiate sequencing, DNA polymerase, fluorescently labelled nucleotides (dNTPs) and modified nucleotides (ddNTPs-ddATP, ddCTP, ddGTP and ddTTP) for termination of DNA strand elongation. Different reactions with DNA template were prepared with all of the requirements except the modified nucleotides. Only one of the ddNTPs was added to each reaction, where its incorporation into the DNA strand inhibited addition of further nucleotides. The ddNTPs do not have a 3'-OH, therefore no further phosphodiester bonds are possible after addition of a ddNTP. Synthesised fragments were denatured and resolved to one base size by gel electrophoresis. The DNA was visualised by UV light and the sequence was read off the gel image (Sanger *et al.*, 1977). Sequence results were analysed using Chromas version 2.01, 2001 and DNAMAN program for Windows, version 5.2.9 (Lynnon BioSoft, © 1994-1999).

Appendix B

Vectors used

i. pUC19 vector

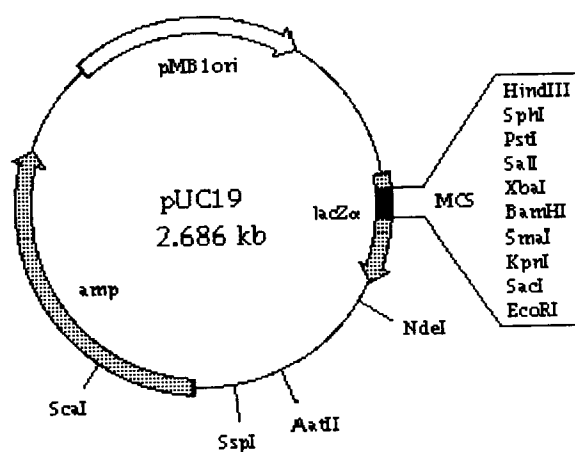


Figure 1: pUC19 cloning vector (www.bioron.net/pUC19-dna.html)

ii. M13 primers used for sequencing

M13 Forward primer 5'-GTTGTAAAACGACGGCT-3'

M13 reverse primer 5'-CAGAAgACAGCTATGAC-3'

iii. pGEM-T Easy vector

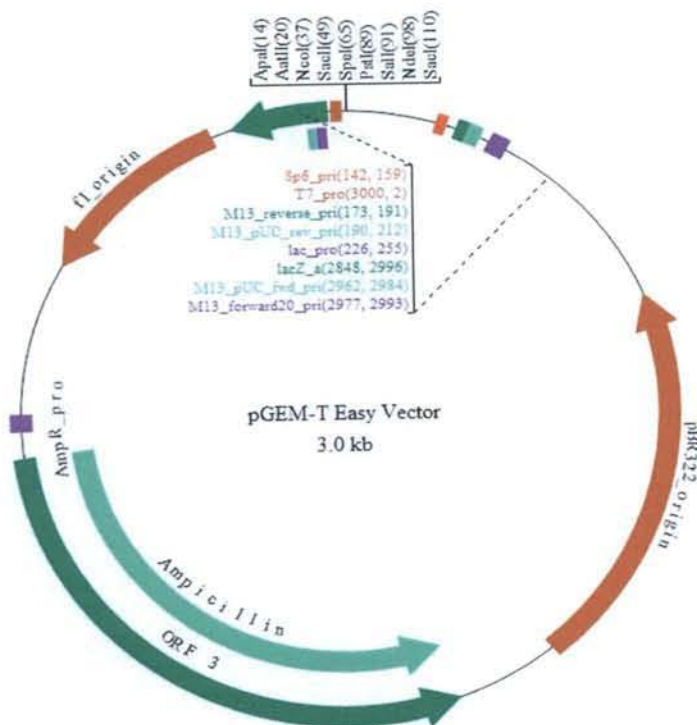


Figure 3: pGEM-T Easy cloning vector (www.biovisual.com/...pGEM-T_Easy_vector.htm)

Appendix C

i. Table 1. Primer sequences for primer-walking of HPV 39 genome

Clone	Primer	Primer sequence (5'-3')	Bases (bp)	Position
15.2	15.2F	TCATCCTGTCCAGGTGCACTGC	24	7593-7614
	15.2.1F	CGGTTGACCTTGTATGTCACGAGC	24	659-682
	15.2.2F	CACAATTTAAAGAAACATATGG	22	1544-1565
	15.2.3F	ACCACTTGCAGATGCAAACTAGC	24	2451-2474
	15.2.4F	TCGACAGTGTGCAGTCACAGAGC	23	3544-3566
	39R	GGTTGGGCATACATACCTAT	20	7807-7826
	15.2.5F	TGTTGTAGATGTGTCTCC	18	4489-4506
	15.2.6F	TTCTTCAGCATCTACTAAATATGC	24	5404-5427
	15.2.7F	TGCCTTTAGATATTTGTCAATCC	23	6298-6320
15.3	15.3F	CATGTCTTTACCTTAGGTTAC	22	7521-7547
	15.3.1F	TACAATGAAATACAGCCGGTTGAC	24	643-666
	15.2.2F	CACAATTTAAAGAAACATATGG	22	1544-1565
	15.2.3F	ACCACTTGCAGATGCAAACTAGC	24	2451-2474
	15.3.4F	TCGACAGTGTGCAGTCACAGAGC	23	3544-3566
	39R	GGTTGGGCATACATACCTAT	20	7807-7826
	15.3.5F	ATGTGTCTCCTGCACGTCCACC	22	4497-4518
	15.3.6F	TTCTTCAGCATCTACTAAATATGC	24	5404-5427
	15.3.7F	GATGTGTATGGGGACAGTATGTTC	24	6357-6380

REFERENCES

1. **Ahmadian, A., and Ehn, E., Hober, S.** 2006. Pyrosequencing: History, biochemistry and future. *Clinica Chimica Acta* **363**:83-94.
2. **Alencar, T. R., D. M. Cerqueira, M. R. Cruz, P. S. Wyant, E. D. Ramalho, and C. R. F. Martins.** 2007. New HPV-16 European and non-European variants in Central Brazil. *Virus Genes* **35**:1-4.
3. **Allan, B., D. J. Marais, M. Hoffman, S. Shapiro, and A. L. Williamson.** 2008. Cervical Human Papillomavirus (HPV) Infection in South African Women: Implications for HPV Screening and Vaccine Strategies. *J. Clin. Microbiol.* **46**:740-742.
4. **Angulo, M., and A. Carvajal-Rodriguez.** 2007. Evidence of recombination within human alpha-papillomavirus. *Viol. J.* **4**:33-40.
5. **Antonsson, A., and B. G. Hansson.** 2002. Healthy skin of many animal species harbors papillomaviruses which are closely related to their human counterparts. *J. Virol.* **76**:12537-12542.
6. **Banura, C., S. Franceschi, L. J. Doorn, A. Arslan, F. Wabwire-Mangen, E. K. Mbidde, W. Quint, and E. Weiderpass.** 2008. Infection with human papillomavirus and HIV among young women in Kampala, Uganda. *J. Infect. Dis.* **197**:555-562.
7. **Barnabas, R., P. Laukkanen, and P. Koskela.** 2006. The epidemiology of HPV16 and cervical cancer in Finland and the potential of vaccination: mathematical modelling analyses. *PLoS Med* **3**:624-632.

8. **Baseman, J. G., and L. A. Koutsky.** 2005. The epidemiology of human papillomavirus infections. *Journal of Clinical Virology* **32**:16-24.
9. **Bentley, D. R.** 2006. Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* **16**:545-552.
10. **Bernard, H. U.** 2005. The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. *Journal of Clinical Virology* **32**:1-6.
11. **Bernard, H. U., I. E. Calleja-Macias, and S. T. Dunn.** 2006. Genome variation of human papillomavirus types: phylogenetic and medical implications. *International Journal of Cancer* **118**:1071-1076.
12. **Bernard, H. U.** 2002. Gene expression of genital human papillomaviruses and considerations on potential antiviral approaches. *Antivir Ther.* **7**:219-237.
13. **Berumen, J., R. M. Ordonez, E. Lazcano, J. Salmeron, S. C. Galvan, R. A. Estrada, E. Yunes, A. Garcia-Carranca, G. Gonzalez-Lira, and A. Madrigal-de la Campa.** 2001. Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. *jnci* **93**:1325-1330.
14. **Bioron.** pUC19 cloning vector. [www.bioron.net/pUC19-dna.html].
15. **Biovisual.** pGEM-T Easy cloning vector. [www.biovisual.com/...pGEM-T_Easy_vector.htm].
16. **Blossom, D. B., R. H. Beigi, J. J. Farrell, W. Mackay, B. Qadadri, D. R. Brown, S. Rwambuya, C. J. Walker, F. S. Kambugu, and F. W. Abdul-Karim.** 2007. Human

papillomavirus genotypes associated with cervical cytologic abnormalities and HIV infection in Ugandan women. *J. Med. Virol.* **79**:758-765.

17. **Boller, K., H. Konig, M. Sauter, N. Mueller-Lantzsch, R. Lower, J. Lower, and R. Kurth.** 1993. Evidence that HERV-K is the endogenous retrovirus sequence that codes for the human teratocarcinoma-derived retrovirus HTDV. *Virology* **196**:349-353.

18. **Brown, C., A. M. Kowalczyk, E. R. Taylor, I. M. Morgan, and K. Gaston.** 2008. P53 represses human papillomavirus type 16 DNA replication via the viral E2 protein. *Virol. J.* **5**:5-13.

19. **Burk, R. D., M. Terai, P. E. Gravitt, L. A. Brinton, R. J. Kurman, W. A. Barnes, M. D. Greenberg, O. C. Hadjimichael, L. Fu, and L. McGowan.** 2003. Distribution of human papillomavirus types 16 and 18 variants in squamous cell carcinomas and adenocarcinomas of the cervix. *Cancer Res.* **63**:7215-7220.

20. **Castellsague, X., and N. Munoz.** 2003. Chapter 3: Cofactors in human papillomavirus carcinogenesis-role of parity, oral contraceptives, and tobacco smoking. *Journal of the National Cancer Institute Monographs* **31**:20-28.

21. **Castle, P. E., C. Porras, W. G. Quint, A. C. Rodriguez, M. Schiffman, P. E. Gravitt, P. Gonzalez, H. A. Katki, S. Silva, E. Freer, L. J. Van Doorn, S. Jimenez, R. Herrero, and A. Hildesheim.** 2008. Comparison of Two PCR-Based Human Papillomavirus Genotyping Methods. *J. Clin. Microbiol.* **46**:3437-3445.

22. **Chan, W. K., G. Klock, and H. U. Bernard.** 1989. Progesterone and glucocorticoid response elements occur in the long control regions of several human papillomaviruses involved in anogenital neoplasia. *J. Virol.* **63**:3261-3269.

23. **Chirenje, Z. M.** 2005. HIV and cancer of the cervix. *Best Practice & Research Clinical Obstetrics & Gynaecology* **19**:269-276.
24. **Clavel, C., M. Masure, I. Putaud, K. Thomas, J. P. Bory, R. Gabriel, C. Quereux, and P. Birembaut.** 1998. Hybrid capture II, a new sensitive test for human papillomavirus detection. Comparison with hybrid capture I and PCR results in cervical lesions. *Br. Med. J.* **51**:737-740.
25. **Clifford, G., S. Franceschi, M. Diaz, N. Muñoz, and L. L. Villa.** 2006a. Chapter 3: HPV type-distribution in women with and without cervical neoplastic diseases. *Vaccine* **24**:26-34.
26. **Clifford, G. M., S. Gallus, R. Herrero, N. Munoz, P. J. F. Snijders, S. Vaccarella, P. T. H. Anh, C. Ferreccio, N. T. Hieu, and E. Matos.** 2005. Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. *The Lancet* **366**:991-998.
27. **Clifford, G. M., M. A. G. Gonçalves, and S. Franceschi.** 2006b. Human papillomavirus types among women infected with HIV: a meta-analysis. *AIDS* **20**:2337-2344.
28. **Coutlee, F., P. Gravitt, J. Kornegay, C. Hankins, H. Richardson, N. Lapointe, H. Voyer, and E. Franco.** 2002. Use of PGMY primers in L1 consensus PCR improves detection of human papillomavirus DNA in genital samples. *J. Clin. Microbiol.* **40**:902-907.

29. **Cox, J. T.** 2006. Human Papillomavirus Testing in Primary Cervical Screening and Abnormal Papanicolaou Management. *Obstet. Gynecol. Surv.* **61**:S15-S25.
30. **de Cremoux, P., J. Coste, X. Sastre-Garau, M. Thioux, C. Bouillac, S. Labbé, I. Cartier, M. Ziol, A. Dosda, and C. Le Galès.** 2003. Efficiency of the Hybrid Capture 2 HPV DNA test in cervical cancer screening. *Am. J. Clin. Pathol.* **120**:492-499.
31. **de Roda Husman, A. M., J. M. Walboomers, A. J. Van den Brule, C. J. Meijer, and P. J. Snijders.** 1995. The use of general primers GP5 and GP6 elongated at their 3'ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. *J. Gen. Virol.* **76**:1057-1062.
32. **de Villiers, E. M.** 1989. Heterogeneity of the human papillomavirus group. *J. Virol.* **63**:4898-4903.
33. **de Villiers, E. M., C. Fauquet, T. R. Broker, H. U. Bernard, and H. zur Hausen.** 2004. Classification of papillomaviruses. *Virology* **324**:17-27.
34. **del Mar Pena, L. M., and L. A. Laimins.** 2002. Regulation of human papillomavirus gene expression in the vegetative life cycle, p. 31-52. *In* D. J. McCance (ed.), *Human papillomaviruses: Perspectives in medical virology*, First ed. Elsevier, Amsterdam, The Netherlands.
35. **Demidov, V. V.** 2005. Rolling-circle amplification (RCA), p. 1175-1179. *In* Fuchs, J., Podda, M. (ed.), *Encyclopedia of Medical Genomics and Proteomics*. Dekker, M., New York.
36. **Doorbar, J.** 2006. Molecular biology of human papillomavirus infection and cervical cancer. *Clin. Sci.* **110**:525-541.

37. **Doorbar, J., and J. C. Sterling.** 2001. The biology of human papillomaviruses, p. 10-23. *In* J. C. Sterling and S. K. Tryling (ed.), Human papillomaviruses: Clinical and scientific advances. Anorl, Great Britain.
38. **Droege, M., and B. Hill.** 2008. The Genome Sequencer FLX™ System—Longer reads, more applications, straight forward bioinformatics and more complete data sets. *J. Biotechnol.* **136**:3-10.
39. **Finegold, S. M.** 1995. Anaerobic infections in humans: an overview. *Anaerobe* **1**:3-9.
40. **Firnhaber, C., K. Zungu, S. Levin, P. Michelow, L. J. Montaner, P. McPhail, A. Williamson, B. R. Allan, C. Van der Horst, and A. Rinas.** 2009. Diverse and high prevalence of human papillomavirus associated with a significant high rate of cervical dysplasia in human immunodeficiency virus-infected women in Johannesburg, South Africa. *Acta Cytol.* **53**:10-17.
41. **Flores, E. R., and P. F. Lambert.** 1997. Evidence for a switch in the mode of human papillomavirus type 16 DNA replication during the viral life cycle. *The Journal of Virology* **71**:7167-7179.
42. **Frazer, I. H.** 2009. Interaction of human papillomaviruses with the host immune system: A well evolved relationship. *Virology* **384**:410-414.
43. **Frisch, M., R. J. Biggar, and J. J. Goedert.** 2000. Human papillomavirus-associated cancers in patients with human immunodeficiency virus infection and acquired immunodeficiency syndrome. *Journal of the National Cancer Institute* **92**:1500-1510.

44. **Gharizadeh, B., M. Kalantari, C. A. Garcia, B. Johansson, and P. Nyrén.** 2001. Typing of human papillomavirus by pyrosequencing. *Laboratory Investigation* **81**:673-679.
45. **Gharizadeh, B., M. Oggionni, B. Zheng, E. Akom, N. Pourmand, A. Ahmadian, K. L. Wallin, and P. Nyren.** 2005. Type-specific multiple sequencing primers: A novel strategy for reliable and rapid genotyping of Human papillomaviruses by pyrosequencing technology. *Journal of Molecular Diagnostics* **7**:198-205.
46. **Gravitt, P. E., C. L. Peyton, T. Q. Alessi, C. M. Wheeler, F. Coutlee, A. Hildesheim, M. H. Schiffman, D. R. Scott, and R. J. Apple.** 2000. Improved amplification of genital human papillomaviruses. *J. Clin. Microbiol.* **38**:357-361.
47. **Gravitt, P. E., C. L. Peyton, R. J. Apple, and C. M. Wheeler.** 1998. Genotyping of 27 human papillomavirus types by using L1 consensus PCR products by a single-hybridization, reverse line blot detection method. *J. Clin. Microbiol.* **36**:3020-3027.
48. **Harwood, A. A., and C. A. Proby.** 2001. Human papillomavirus and immunosuppression, p. 102-117. *In* J. C. Sterling and S. K. Tryg (ed.), *Anorl*, Great Britain.
49. **Hawes, S. E., C. W. Critchlow, M. A. Faye Niang, M. B. Diouf, A. Diop, P. Toure, A. Aziz Kasse, B. Dembele, P. Salif Sow, A. M. Coll-Seck, J. M. Kuypers, and N. B. Kiviat.** 2003. Increased risk of high-grade cervical squamous intraepithelial lesions and invasive cervical cancer among African women with human immunodeficiency virus type 1 and 2 infections. *J. Infect. Dis.* **188**:555-563.

50. **Hazard, K., K. Andersson, J. Dillner, and O. Forslund.** 2007. Human papillomavirus subtypes are not uncommon. *Virology* **362**:6-9.
51. **Herrero, R.** 2009. Human Papillomavirus (HPV) Vaccines: Limited cross-protection against additional HPV types. *J. Infect. Dis.* **199**:919-922.
52. **Hesselink, A. T., M. van Ham, D. A. M. Heideman, Z. M. A. Groothuismink, L. Rozendaal, J. Berkhof, F. J. van Kemenade, L. Massuger, W. J. G. Melchers, and C. Meijer.** 2008. Comparison of GP5+/6+ -PCR and SPF10-Line blot assays for detection of high-risk Human papillomavirus in samples from women with normal cytology results who develop grade 3 cervical intraepithelial neoplasia. *J. Clin. Microbiol.* **46**:3215-3221.
53. **Hoory, T., A. Monie, P. Gravitt, and T. C. Wu.** 2008. Molecular epidemiology of human papillomavirus. *J. Formosan Med. Assoc.* **107**:198-217.
54. **Howley, P. M.** 1996. Papillomavirinae: The viruses and their replication, p. 2045-2076. *In* B. N. Fields, D. M. Knipe and P. M. Howley (ed.), *Fields Virology*, 3rd ed. Lippincott- Raven, Philadelphia.
55. **Howley, P. M., and D. R. Lowy.** 2007. Papillomaviruses, p. 2299-2354. *In* D. M. Knipe and P. M. Howley (ed.), *Fields Virology*, 5th ed. Lippincott-Raven, Philadelphia.
56. **Huang, S. L., A. Chao, S. Hsueh, F. Y. Chao, C. C. Huang, J. E. Yang, C. Y. Lin, C. C. Yan, H. H. Chou, and K. G. Huang.** 2006. Comparison between the Hybrid Capture II Test and an SPF1/GP6 PCR-based assay for detection of human papillomavirus DNA in cervical swab samples. *J. Clin. Microbiol.* **44**:1733-1739.

57. **International Agency for Research on Cancer.** 2007. IARC monographs on the evaluation of carcinogenic risks to humans. Human Papillomaviruses. WHO Press, Switzerland.
58. **Jacobs, M. V., P. J. Snijders, A. J. Van den Brule, T. J. Helmerhorst, C. J. Meijer, and J. M. Walboomers.** 1997. A general primer GP5+ /GP6 (+)-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings. *J. Clin. Microbiol.* **35**:791-795.
59. **Jeon, S., B. L. Allen-Hoffmann, and P. F. Lambert.** 1995. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J. Virol.* **69**:2989-2997.
60. **Kadaja, M., T. Silla, E. Ustav, and M. Ustav.** 2009. Papillomavirus DNA replication—From initiation to genomic instability. *Virology* **384**:360-368.
61. **Kahn, T., E. Schwarz, and H. Z. Hausen.** 1986. Molecular cloning and characterization of the DNA of a new human papillomavirus (HPV 30) from a laryngeal carcinoma. *International Journal of Cancer* **37**:61-65.
62. **Kämmer, C., M. Tommasino, S. Syrjänen, H. Delius, U. Hebling, U. Warthorst, H. Pfister, and I. Zehbe.** 2002. Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. *Br. J. Cancer* **86**:269-273.
63. **Kämmer, C., U. Warthorst, N. Torrez-Martinez, C. M. Wheeler, and H. Pfister.** 2000. Sequence analysis of the long control region of human papillomavirus type 16

variants and functional consequences for P97 promoter activity. *J. Gen. Virol.* **81**:1975-1981.

64. **Kjaer, S. K., A. J. C. van den Brule, J. E. Bock, P. A. Poll, G. Engholm, M. E. Sherman, J. M. M. Walboomers, and C. J. L. M. Meijer.** 1996. Human papillomavirus-the most significant risk determinant of cervical intraepithelial neoplasia. *International Journal of Cancer* **65**:601-606.

65. **Kleter, B., L. J. Van Doorn, L. Schrauwen, A. Molijn, S. Sastrowijoto, J. ter Schegget, J. Lindeman, B. ter Harmsel, M. Burger, and W. Quint.** 1999. Development and clinical evaluation of a highly sensitive PCR-reverse hybridization line probe assay for detection and identification of anogenital human papillomavirus. *J. Clin. Microbiol.* **37**:2508-2517.

66. **Kleter, B., L. J. van Doorn, J. ter Schegget, L. Schrauwen, K. van Krimpen, M. Burger, B. ter Harmsel, and W. Quint.** 1998. Novel short-fragment PCR assay for highly sensitive broad-spectrum detection of anogenital human papillomaviruses. *Am. J. Pathol.* **153**:1731-1739.

67. **Kurvinen, K., M. Yliskoski, S. Saarikoski, K. Syrjänen, and S. Syrjänen.** 2000. Variants of the long control region of human papillomavirus type 16. *Eur. J. Cancer* **36**:1402-1410.

68. **Leman, J. A and Benton E.** 2000. Verrucas: Guidelines for management. *Am. J. Clin Dermatology.* **1**:143-149.

69. **Levi, J. E., B. Kleter, W. G. Quint, M. C. Fink, C. L. Canto, R. Matsubara, I. Linhares, A. Segurado, B. Vanderborght, J. E. Neto, and L. J. Van Doorn.** 2002.

High prevalence of human papillomavirus (HPV) infections and high frequency of multiple HPV genotypes in human immunodeficiency virus-infected women in Brazil. *J. Clin. Microbiol.* **40**:3341-3345.

70. **Lin, C. Y., A. Chao, Y. C. Yang, H. H. Chou, C. M. Ho, R. W. Lin, T. C. Chang, J. Y. Chiou, F. Y. Chao, and K. L. Wang.** 2008. Human papillomavirus typing with a polymerase chain reaction-based genotyping array compared with type-specific PCR. *Journal of Clinical Virology* **42**:361-367.

71. **Liu, J. S., and T. Melendy.** 2002. Human papillomavirus DNA replication, p. 53-70. *In* D. J. McCance (ed.), *Human papillomaviruses: Perspectives in medical virology*, First ed. Elsevier, Amsterdam, The Netherlands.

72. **Lizano, M., J. Berumen, and A. García-Carrancá.** 2009. HPV-related Carcinogenesis: Basic Concepts, Viral Types and Variants. *Arch. Med. Res.* **40**:428-434.

73. **Longuet, M., P. Cassonnet, and G. Orth.** 1996. A novel genital human papillomavirus (HPV), HPV type 74, found in immunosuppressed patients. *J. Clin. Microbiol.* **34**:1859-1862.

74. **Longworth, M. S., and L. A. Laimins.** 2004. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiology and Molecular Biology Reviews* **68**:362-372.

75. **Luque, A. E., M. Jabeen, S. Messing, C. A. Lane, L. M. Demeter, R. C. Rose, and R. C. Reichman.** 2006. Prevalence of human papillomavirus genotypes and related abnormalities of cervical cytological results among HIV-1-infected women in Rochester, New York. *J. Infect. Dis.* **194**:428-434.

76. **Luque, A. E., L. M. Demeter, and R. C. Reichman.** 1999. Association of human papillomavirus infection and disease with magnitude of human immunodeficiency virus type 1 (HIV-1) RNA plasma level among women with HIV-1 infection. *J. Infect. Dis.* **179**:1405-1409.
77. **Manos, M. M., Y. Ting, D. K. Wright, A. J. Lewis, T. R. Broker, and S. M. Wolinsky.** 1989. Use of PCR amplification for the detection of genital HPV. *Cancer Cells* **7**:209-214.
78. **Marais, D. J., J. A. S. Passmore, L. Denny, C. Sampson, B. R. Allan, and A. L. Williamson.** 2008. Cervical and oral human papillomavirus types in HIV-1 positive and negative women with cervical disease in South Africa. *J. Med. Virol.* **80**:953-959.
79. **Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bembien, J. Berka, M. S. Braverman, Y. J. Chen, and Z. Chen.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**:376-380.
80. **McLaughlin-Drubin, M. E., and K. Münger.** 2009. Oncogenic activities of human papillomaviruses. *Virus Res.* **143**:195-208.
81. **Moodley, J. R., D. Constant, M. Hoffman, A. Salimo, B. Allan, E. Rybicki, I. Hitzeroth, and A. L. Williamson.** 2009. Human papillomavirus prevalence, viral load and pre-cancerous lesions of the cervix in women initiating highly active antiretroviral therapy in South Africa: a cross-sectional study. *BMC Cancer* **9**:275-282.
82. **Moodley, J. R., M. Hoffman, H. Carrara, B. R. Allan, D. D. Cooper, L. Rosenberg, L. E. Denny, S. Shapiro, and A. L. Williamson.** 2006. HIV and pre-

neoplastic and neoplastic lesions of the cervix in South Africa: a case-control study. *BMC Cancer* **6**:135-140.

83. **Mqoqi, N., P. Kellet, F. Sitas, and M. Jula.** 2004. Cancer in South Africa, 1998-1999. National Cancer Registry of South Africa, National Health Laboratory Service, South Africa.

84. **Münger, K., and P. M. Howley.** 2002. Human papillomavirus immortalization and transformation functions. *Virus Res.* **89**:213-228.

85. **Munoz, N., F. X. Bosch, S. de Sanjose, R. Herrero, X. Castellsague, K. V. Shah, P. J. F. Snijders, and C. J. L. M. Meijer.** 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.* **348**:518-527.

86. **Muñoz, N., X. Castellsagué, A. B. de González, and L. Gissmann.** 2006. Chapter 1: HPV in the etiology of human cancer. *Vaccine* **24**:1-10.

87. **National Center for Biotechnology Information.** 2009. September 2009. NCBI [www.ncbi.nlm.nih.gov].

88. **Naucler, P., W. Ryd, S. Tornberg, A. Strand, G. Wadell, K. Elfgrén, T. Radberg, B. Strander, B. Johansson, and O. Forslund.** 2007. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N. Engl. J. Med.* **357**:1589.

89. **Palefsky, J.** 2006. Biology of HPV in HIV infection. *Adv. Dent. Res.* **19**:99-105.

90. **Palefsky, J. M., and E. A. Holly.** 2003. Chapter 6: Immunosuppression and co-infection with HIV. *Journal of the National Cancer Institute Monographs* **2003**:41-46.

91. **Palefsky, J. M., E. A. Holly, J. Gonzales, J. Berline, D. K. Ahn, and J. S. Greenspan.** 1991. Detection of human papillomavirus DNA in anal intraepithelial neoplasia and anal cancer. *Cancer Res.* **51**:1014-1019.
92. **Pande, S., N. Jain, B. K. Prusty, S. Bhambhani, S. Gupta, R. Sharma, S. Batra, and B. C. Das.** 2008. Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and Long Control Region in biopsy samples from cervical cancer patients in North India. *Journal of Clinical Virology* **46**:1060-1066.
93. **Peterlin, B. M., and D. Trono.** 2003. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. *Nature Reviews Immunology* **3**:97-107.
94. **Poljak, M., A. Brencic, K. Seme, A. Vince, and I. J. Marin.** 1999. Comparative evaluation of first- and second-generation Digene Hybrid Capture assays for detection of human papillomaviruses associated with high or intermediate risk for cervical cancer. *J. Clin. Microbiol.* **37**:796-797.
95. **Prado, J. C., I. E. Calleja-Macias, H. U. Bernard, M. Kalantari, S. A. Macay, B. Allan, A. L. Williamson, L. P. Chung, R. J. Collins, and R. E. Zuna.** 2005. Worldwide genomic diversity of the human papillomaviruses-53, 56, and 66, a group of high-risk HPVs unrelated to HPV-16 and HPV-18. *Virology* **340**:95-104.
96. **Rector, A., R. Tachezy, K. Van Doorslaer, T. MacNamara, R. D. Burk, J. P. Sundberg, and M. Van Ranst.** 2005. Isolation and cloning of a papillomavirus from a North American porcupine by using multiply primed rolling-circle amplification: the *Erethizon dorsatum* papillomavirus type 1. *Virology* **331**:449-456.

97. **Rector, A., R. Tachezy, and M. Van Ranst.** 2004. A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. *J. Virol.* **78**:4993-4998.
98. **Ronaghi, M.** 2001. Pyrosequencing sheds light on DNA sequencing. *Genome Res.* **11**:3-11.
99. **Safaeian, M., R. Herrero, A. Hildesheim, W. Quint, E. Freer, L. J. Van Doorn, C. Porras, S. Silva, P. González, and M. C. Bratti.** 2007. Comparison of the SPF₁₀-LiPA System to the Hybrid Capture 2 assay for detection of carcinogenic Human papillomavirus genotypes among 5,683 young women in Guanacaste, Costa Rica. *J. Clin. Microbiol.* **45**:1447-1454.
100. **Sahasrabudde, V. V., M. H. Mwanahamuntu, S. H. Vermund, W. K. Huh, M. D. Lyon, J. S. A. Stringer, and G. P. Parham.** 2007. Prevalence and distribution of HPV genotypes among HIV-infected women in Zambia. *Br. J. Cancer* **96**:1480-1483.
101. **Said, H. M., K. Ahmed, R. Burnett, B. R. Allan, A. L. Williamson, and A. A. Hoosen.** 2009. HPV genotypes in women with squamous intraepithelial lesions and normal cervixes participating in a community-based microbicide study in Pretoria, South Africa. *Journal of Clinical Virology* **44**:318-321.
102. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406-425.
103. **Sambrook, J., and E. F. Fritsch.** 1989. *Molecular cloning: a laboratory manual.* Cold Spring Harbour Laboratory Press, New York.

104. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**:5463-5467.
105. **Schiffman, M., G. Clifford, and F. M. Buonaguro.** 2009. Classification of weakly carcinogenic human papillomavirus types: addressing the limits of epidemiology at the borderline. *Infectious Agents and Cancer* **4**:8-15.
106. **Schiffman, M., R. Herrero, R. DeSalle, A. Hildesheim, S. Wacholder, A. Cecilia Rodriguez, M. C. Bratti, M. E. Sherman, J. Morales, and D. Guillen.** 2005. The carcinogenicity of human papillomavirus types reflects viral evolution. *Virology* **337**:76-84.
107. **Schiffman, M., and S. K. Kjaer.** 2003. Chapter 2: Natural history of anogenital human papillomavirus infection and neoplasia. *Journal of the National Cancer Institute Monographs* **2003**:14-19.
108. **Schmitt, M., B. Dondog, T. Waterboer, and M. Pawlita.** 2008. Homogeneous amplification of genital human alpha papillomaviruses by PCR using novel broad-spectrum GP5 and GP6 primers. *J. Clin. Microbiol.* **46**:1050-1059.
109. **Seme, K., K. Fujs, B. J. Kocjan, and M. Poljak.** 2006. Resolving repeatedly borderline results of Hybrid Capture 2 HPV DNA Test using polymerase chain reaction and genotyping. *J. Virol. Methods* **134**:252-256.
110. **Shah, K. V., and P. M. Howley.** 1996. Papillomaviruses, p. 2077-2105. *In* B. N. Fields, D. M. Knipe and P. M. Howley (ed.), *Fields Virology*, 3rd ed. Lippincott-Raven, Philadelphia.

111. **Shapiro, S., L. Rosenberg, M. Hoffman, J. P. Kelly, D. D. Cooper, H. Carrara, L. E. Denny, G. du Toit, B. R. Allan, and I. A. Stander.** 2003. Risk of invasive cancer of the cervix in relation to the use of injectable progestogen contraceptives and combined estrogen/progestogen oral contraceptives (South Africa). *Cancer Causes and Control* **14**:485-495.
112. **Singh, D. K., K. Anastos, D. R. Hoover, R. D. Burk, Q. Shi, L. Ngendahayo, E. Mutimura, A. Cajigas, V. Bigirimani, and X. Cai.** 2009. Human Papillomavirus Infection and Cervical Cytology in HIV-Infected and HIV-Uninfected Rwandan Women. *J. Infect. Dis.* **199**:1851-1861.
113. **Soderlund-Strand, A., P. Rymark, P. Andersson, J. Dillner, and L. Dillner.** 2005. Comparison between the Hybrid Capture II test and a PCR-based human papillomavirus detection method for diagnosis and posttreatment follow-up of cervical intraepithelial neoplasia. *J. Clin. Microbiol.* **43**:3260-3266.
114. **Stanley, M.** 2003. Genital Human papillomavirus Infections--Current and prospective therapies. *JNCI Cancer Spectrum* **31**:117-124.
115. **Stewart, A. C., A. M. Eriksson, M. M. Manos, N. Munoz, F. X. Bosch, J. Peto, and C. M. Wheeler.** 1996. Intratype variation in 12 human papillomavirus types: a worldwide perspective. *J. Virol.* **70**:3127-3136.
116. **Stunkel, W., Z. Huang, S. H. Tan, M. J. O'Connor, and H. U. Bernard.** 2000. Nuclear matrix attachment regions of human papillomavirus type 16 repress or activate the E6 promoter, depending on the physical state of the viral DNA. *J. Virol.* **74**:2489-2501.

117. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**:1596-1599.
118. **Tamura, K., M. Nei, and S. Kumar.** 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences* **101**:11030-11035.
119. **Tanzi, E., A. Amendola, S. Bianchi, M. M. Fasolo, R. Beretta, E. Pariani, A. Zappa, E. Frati, and G. Orlando.** 2009. Human papillomavirus genotypes and phylogenetic analysis of HPV-16 variants in HIV-1 infected subjects in Italy. *Vaccine* **27**:17-23.
120. **Thomison, J., L. K. Thomas, and K. R. Shroyer.** 2008. Human papillomavirus: molecular and cytologic/histologic aspects related to cervical intraepithelial neoplasia and carcinoma. *Hum. Pathol.* **39**:154-166.
121. **Tjalma, W. A. A., T. R. Van Waes, L. E. M. Van den Eeden, and J. Bogers.** 2005. Role of human papillomavirus in the carcinogenesis of squamous cell carcinoma and adenocarcinoma of the cervix. *Best Practice & Research Clinical Obstetrics & Gynaecology* **19**:469-483.
122. **Tornesello, M. L., F. M. Buonaguro, L. Buonaguro, I. Salatiello, E. Beth-Giraldo, and G. Giraldo.** 2000. Identification and functional analysis of sequence rearrangements in the long control region of human papillomavirus type 16 Af-1 variants isolated from Ugandan penile carcinomas. *J. Gen. Virol.* **81**:2969-2982.

123. **Tornesello, M. L., M. L. Duraturo, P. Giorgi-Rossi, M. Sansone, R. Piccoli, L. Buonaguro, and F. M. Buonaguro.** 2008. Human papillomavirus (HPV) genotypes and HPV16 variants in human immunodeficiency virus-positive Italian women. *J. Gen. Virol.* **89**:1380-1389.
124. **Tornesello, M. L., M. L. Duraturo, I. Salatiello, L. Buonaguro, S. Losito, G. Botti, G. Stellato, S. Greggi, R. Piccoli, and S. Pilotti.** 2004. Analysis of human papillomavirus type-16 variants in Italian women with cervical intraepithelial neoplasia and cervical cancer. *J. Med. Virol.* **74**:117-126.
125. **Trottier, H., and E. L. Franco.** 2006. The epidemiology of genital human papillomavirus infection. *Vaccine* **24**:4-15.
126. **Tu, J. J., L. Kuhn, L. Denny, K. J. Beattie, A. Lorincz, and T. C. Wright Jr.** 2006. Molecular variants of human papillomavirus type 16 and risk for cervical neoplasia in South Africa. *Int J Gynecol Cancer* **16**:736-742.
127. **UNAIDS.** 2008. Global report on HIV data.
[www.unaids.org/en/Knowledge/HIVData/Globalreport/2008].
128. **van den Brule, A. J. C., R. Pol, N. Fransen-Daalmeijer, L. M. Schouls, C. J. L. M. Meijer, and P. J. F. Snijders.** 2002. GP5+/6+ PCR followed by reverse line blot analysis enables rapid and high-throughput identification of human papillomavirus genotypes. *J. Clin. Microbiol.* **40**:779-787.
129. **Varsani, A., E. van der Walt, L. Heath, E. P. Rybicki, A. L. Williamson, and D. P. Martin.** 2006. Evidence of ancient papillomavirus recombination. *J. Gen. Virol.* **87**:2527-2531.

130. **Vernon, S. D., C. E. Hart, W. C. Reeves, and J. P. Icenogle.** 1993. The HIV-1 tat protein enhances E2-dependent human papillomavirus 16 transcription. *Virus Res.* **27**:133-145.
131. **Vinokurova, S., N. Wentzensen, J. Eienkel, R. Klaes, C. Ziegert, P. Melsheimer, H. Sartor, L. C. Horn, M. Hockel, and M. von Knebel Doeberitz.** 2005. Clonal history of papillomavirus-induced dysplasia in the female lower genital tract. *JNCI Journal of the National Cancer Institute* **97**:1816-1821.
132. **Walboomers, J. M. M., M. V. Jacobs, M. M. Manos, F. X. Bosch, J. A. Kummer, K. V. Shah, P. J. F. Snijders, J. Peto, C. J. L. M. Meijer, and N. Munoz.** 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.* **189**:12-19.
133. **Wexler, H. M.** 2007. Bacteroides: the good, the bad, and the nitty-gritty. *Clin. Microbiol. Rev.* **20**:593-623.
134. **Wheeler, C. M.** 2002. Clinical aspects and epidemiology of HPV infections, p. 1-30. *In* D. J. McCance (ed.), *Human papillomaviruses: Perspectives in medical virology*, First ed. Elsevier Science, Amsterdam, The Netherlands.
135. **WHO/ICO.** 2009. Human papillomavirus and related cancers in world. World Health Organisation.
[www.who.int/hpvcenter/statistics/dynamic/ico/SummaryReportSelect.cfm].
136. **Wicker, T., E. Schlagenhauf, A. Graner, T. J. Close, B. Keller, and N. Stein.** 2006. 454 Sequencing put to the test using the complex genome of barley. *BMC Genomics* **7**:275-285.

137. **Woolford, L., A. Rector, M. Van Ranst, A. Ducki, M. D. Bennett, P. K. Nicholls, K. S. Warren, R. A. Swan, G. E. Wilcox, and A. J. O'Hara.** 2007. A novel virus detected in papillomas and carcinomas of the endangered western barred bandicoot (*Perameles bougainville*) exhibits genomic features of both the Papillomaviridae and Polyomaviridae. *J. Virol.* **81**:13280-13290.
138. **Wright Jr, T. C., L. S. Massad, C. J. Dunton, M. Spitzer, E. J. Wilkinson, and D. Solomon.** 2007. Consensus guidelines for the management of women with abnormal cervical cancer screening tests. *Obstet. Gynecol.* **197**:346-355.
139. **Yamada, T., M. M. Manos, J. Peto, C. E. Greer, N. Munoz, F. X. Bosch, and C. M. Wheeler.** 1997. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *The Journal of Virology* **71**:2463-2472.
140. **zur Hausen, H.** 2000. Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *Journal of the National Cancer Institute* **92**:690-698.